# Computation of Greeks in Financial Markets Driven by Lévy Processes

**by**

**Sindre Froyn**

***THESIS***

*for the degree of*

***MASTER OF SCIENCE***

*(Master i Modellering og Dataanalyse)*

*Faculty of Mathematics and Natural Sciences*
*University of Oslo*

*May 2012*

*Det matematisk- naturvitenskapelige fakultet*
*Universitetet i Oslo*

# Abstract

This thesis is divided into 4 chapters. Chapter 1 gives a brief explanation of what the Greeks are and why they are of interest in applied financial mathematics. There is also a short summary of the first attempts at numerical methods to calculate the Greeks as well as an introduction to Lévy processes.

Chapter 2 starts with some relevant results from Malliavin Calculus and proceeds to derivations of general expressions for the most important Greeks using Malliavin weights. It concludes with a mathematical argument that shows how these weights can be regarded as optimal.

Chapter 3 introduces stochastic volatility models followed by some more detailed analysis of a specific stochastic volatility model called the Barndorff-Nielsen and Shephard model. The technicalities involved in doing the necessary simulations for this model are discussed and implemented in Matlab.

Chapter 4 contains a summary and outlines possible extensions to this thesis.

# Acknowledgements

I would like to thank my supervisor, Frank Proske, for providing me with a very interesting topic for my thesis, as well as for his help and encouragement.

I would like to thank my family and friends for proof reading and helping me find misprints and poor formulations.

# Notation

When a numbered equation is referred to, it will be by chapter and number, e.g (2.5) which will be the fifth equation in chapter 2. When a result (definition, theorem etc.) is referred to, it will be by result type, chapter and number, e.g (D2.5) refers to Definition 2.5, the fifth result in chapter 2. Figures are simply referred to as Figure 2.5.

Sources are referred to by numbers only, e.g [4], with the complete list of sources found in the Bibliography on page 62.

Some norms that will be used without specification:

$$\|g\|^2_{L^2([0,T]^n)} := \int_0^T \cdots \int_0^T g^2(t_1, \dots, t_n) dt_1 \dots dt_n$$
$$(g,h)_{L^2([0,T]^n)} := \int_0^T \cdots \int_0^T g(t_1, \dots, t_n) h(t_1, \dots, t_n) dt_1 \dots dt_n$$

Note that $(g,g)_{L^2([0,T]^n)} = \|g\|^2_{L^2([0,T]^n)}$.

$$\|g\|^2_{L^2(P\times\lambda)} := \mathbb{E}\Big[\int_0^T g^2(t)dt\Big]$$

For the probability space $(\Omega, \mathcal{F}, P)$ the standard $L^2$-norm is:

$$\|X\|^2_{L^2(P)} := \mathbb{E}[|X|^2] = \int_\Omega |X|^2(\omega) dP(\omega)$$

The more general $L^p$-norm is:

$$\|X\|^p_{L^p(P)} := \mathbb{E}[|X|^p] = \int_\Omega |X|^p(\omega) dP(\omega)$$

Notation for the indicator function:

$$\mathbf{1}_A(x) := \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

This thesis discusses two different concepts of *derivatives*; the first type being a financial derivative such as an option or a future, and the second type being the standard mathematical notion. The context makes it clear which type is being used, but the former is usually stressed as a *financial derivative*.

# Contents

# Chapter 1

# Introduction

Any investment in a financial market comes with a certain amount of risk. The value of the investment can be severely reduced if the market moves in an unfavourable way, and in the worst case scenario the investment can become completely worthless. Ever since the modern portfolio theory was introduced in the 1950s, the reduction of risk has been recognised as vital in the management of financial assets. The simplest form of risk reduction, also called *hedging*, is to diversify the investment into assets that tend to move in opposite directions (or more precisely; assets that are *negatively correlated*), such as bonds versus stocks, or stocks in airlines versus the oil industry.

The 1970s saw the introduction of new types of financial assets called derivatives, such as options, which essentially are contracts based on other financial assets like stocks or commodities. Among other things, the introduction of derivatives provided an efficient way to reduce portfolio risk, as well as giving rise to new methods of speculation. Financial derivatives have become immensely popular with various estimates[1] placing the total annual value of the derivatives market in the range of several hundred trillions of US dollars, in many cases even exceeding the value of the markets of the underlying assets!

Given the large amounts of capital involved, hedging away risk associated with a financial derivative becomes of great interest and importance in applied financial mathematics. It turns out that the necessary strategy required to hedge away the risk can be found through a set of quantities known as the *sensitivity parameters*, more commonly referred to as the *Greeks*.

The Greeks are unobservable parameters in the market, so the calculation methods to find them depend completely on the choice of the model for the underlying assets on which the derivative is based. This is just one of many reasons to model financial assets as accurately as possible (in the historical sense), which in turn motivates the introduction of the so called *Lévy processes* when building stochastic processes to model financial assets.

[1]BIS Quarterly Review

The purpose of this thesis is to provide a detailed treatment of the Greeks. Various methods used to calculate the Greeks are discussed and compared, with special emphasis on the method involving Malliavin weights, which will be considered in both the traditional, continuous case as well as in the discontinuous Lévy model case.

There are four chapters, the first chapter introducing some background material for the rest of the paper. Chapter 1 starts off with a proper introduction to the four most important Greeks that will be the focus of this thesis. The first methods of finding numerical approximations to the Greeks are mentioned, and the chapter concludes by formally introducing Lévy processes.

Chapter 2 is the main chapter and starts off by listing some central results and definitions from Malliavin calculus which will be used thereafter. The chapter continues with a thorough discussion of a method of calculating the Greeks by using Malliavin calculus and concludes with an examination of how they can be considered optimal in the minimal variance sense.

Chapter 3 introduces the BNS model and demonstrates how to analytically derive and numerically calculate the Greeks in a Lévy market model. The implementation is done in Matlab.

A short summary as well as possible extensions to this thesis are covered in Chapter 4. A collection of relevant calculations and results, in addition to the source code for the Matlab programs used, can be found in the Appendix.

## 1.1 Greeks: the Sensitivity Parameters

This section gives an introduction to the sensitivity parameters that will be discussed in this thesis. They are introduced within the framework of the Black-Scholes market, but the mathematical definitions of the Greeks carry over to more general settings.

### Delta

The most important Greek is the delta, denoted by the Greek letter $\Delta$, which will be derived in the same fashion as presented in [20]. Under the assumptions of the Black-Scholes market, the stock prices are modelled by a geometric Brownian motion, given by the stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = x, \tag{1.1}$$

where the initial price $x$, the drift $\mu$ and volatility $\sigma$ are assumed to be positive constants, and $W_t$ is the standard Brownian motion (or equivalently the Wiener process).

Given a financial derivative in the form of a call option on some underlying stock $S_t$, the option price is given by $V(t, S_t)$. The owner of the option stands to make a profit if the underlying stock price rises. However, there is also a risk present, as the owner of the option may incur a loss if the stock price falls. By taking advantage of the positive correlation between the call option and the underlying stock $S_t$, it is possible to hedge against the risk by shorting the stock. In a short position the situation is reversed, where a profit is made if the stock falls, and a loss incurred if the stock rises. The amount of stock that must be shorted to maintain a balance between the two financial positions is the $\Delta$.

Introducing the portfolio:

$$\Pi = V(t, S_t) - \Delta S_t, \tag{1.2}$$

which consists of the call option with value $V(t, S_t)$ and a $\Delta$ short position in $S_t$. The infinitesimal change in the portfolio is:

$$d\Pi = dV(t, S_t) - \Delta dS_t. \tag{1.3}$$

Applying Ito's lemma to $dV(t, S_t)$, (details provided in (LA.1) on page 49):

$$dV(t, S_t) = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial x} dS_t + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2} dt.$$

Substituting this equality into equation (1.3):

$$\begin{aligned} d\Pi &= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial x} dS_t + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2} dt - \Delta dS_t \\ &= \Big(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2}\Big) dt + \Big(\frac{\partial V}{\partial x} - \Delta\Big) dS_t \end{aligned}$$

By choosing $\Delta = \frac{\partial V}{\partial x}(t, S_t)$, we eliminate the small fluctuations in the change of the stock price, $S_t$, and achieve the delta neutral position.

$$= \Big(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2}\Big) dt + \cancel{\Big(\frac{\partial V}{\partial x} - \frac{\partial V}{\partial x}\Big)} dS_t.$$

The $\Delta$ is usually referred to as the sensitivity of the option with respect to the stock price $x$, and is a measure of how movements in the stock price affect the option value. It is formally defined as the derivative of the option value with respect to the stock price as seen above: $\Delta := \frac{\partial V}{\partial x}(t, S_t)$.

The $\Delta$ is a time-dependent parameter and by continuously readjusting the shorted position to maintain delta neutrality (a process called delta hedging), it is theoretically possible to eliminate the risk associated with the underlying stock $S_t$.

## Gamma

The ability to perfectly delta hedge is not realistic. In the framework of the Black-Scholes market it is possible as the assumptions allow for owning fractions of stocks, continuous trading, no transaction costs (frictionless market) and there are no restrictions on the amount of available stocks in the market. None of these assumptions apply to reality, so the best practical course of action is to approximate the delta at discrete time points.

To reduce the amount of re-hedging required, the sensitivity of the $\Delta$ with respect to the stock price $x$ will be used. This is the second sensitivity parameter known as the gamma, denoted by $\Gamma$, and is defined as $\Gamma := \frac{\partial}{\partial x}\Delta = \frac{\partial^2 V}{\partial x^2}(t, S_t)$.

The $\Gamma$ is, according to [20], a measure of how often or how much a position must be re-hedged in order to maintain a delta neutral position, so to minimize the amount of necessary re-hedging and the corresponding cost, it is possible to expand the portfolio $\Pi$ from (1.2) with additional options to achieve a $\Gamma$ neutral position, i.e a position where $\Gamma = 0$.

## Vega

The volatility of the model, which is a measure of risk, is the key parameter for the value of the option, and for hedging purposes it is important to know how the stock price is affected by movements in the volatility. This leads to the definition of the third sensitivity parameter: vega, denoted by $\nu$, which is defined as $\nu := \frac{\partial V}{\partial \sigma}$.

For $\Delta$ and $\Gamma$ we find the derivative with respect to an observable variable, namely the stock price $x$, but for $\nu$ we are calculating the derivative with respect to a model parameter.

Vega hedging means including additional options to the portfolio with the goal of achieving $\nu = 0$. The true volatility is an unobservable quantity in the market and in a $\nu$ neutral position the exposure to the volatility has been decreased, making the portfolio more insensitive to volatility fluctuations.

## Rho

The Greek rho, denoted by $\rho$ and most commonly defined as $\rho := \frac{\partial V}{\partial r}$, where $r$ is the risk free interest rate from the Black-Scholes market, is different from the previous Greeks as it is not used to hedge away risk. Instead, $\rho$, which measures how the option value changes when the interest rate does, is in practice primarily used to monitor the portfolio.

A more general characterisation of $\rho$ is by defining it as the derivative with respect to the model drift, i.e $\rho := \frac{\partial V}{\partial \mu}$. We regard $r$ as the drift for the geometric Brownian motion after we have applied Girsanov's Theorem, so the former definition is in a way a special case of the latter one.

### Additional Greeks

There are a number of Greeks that will not be discussed, of which the most prominent one is the Theta, defined as $\Theta := \frac{\partial V}{\partial t}(t, S_t)$, which is the sensitivity of the option value with respect to the time left before the option expires.

There are also many more higher order Greeks such as the Speed, Vanna, Vomma, Ultima, etc, but they are not as commonly used in practice. For this paper, we restrict our attention to the four Greeks introduced above.

## 1.2 Methods of Numerical Calculations

In the Black-Scholes market it is possible to calculate the derivatives of the option value and get explicit expressions for the Greeks. In general this can't be done and in most market models the Greeks must be calculated by numerical approximations. Two such methods are briefly discussed in this section.

### The Finite Difference Method

The finite difference method serves as one of the simplest types of Monte Carlo simulation techniques that can be applied to calculating the Greeks numerically. Based on [12] we give a description of how the finite difference method is used to calculate $\Delta$ and $\Gamma$, and as an illustration we will be using geometric Brownian motion to model the stock price $S_t$.

The solution to the dynamics in (1.1) is showed in (LA.2) on page 50 to be (with $S_0 = x$, and using $\mu = r$, so we are working under the risk neutral probability measure):

$$S_t = x \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}.$$

Calculating the Greeks using the finite difference method requires calculations of the option price for different starting values $x$ for the stock $S_t$. In the following the other model parameters are assumed to be kept constant.

By dividing the interval $[0, T]$ (assuming the option expires at time $T$) into $n$ equal parts of length $\Delta t = \frac{T}{n}$, a simulated path of $S_t$ is given by:

$$\mathbf{S} = (S_0, S_1, S_2, \dots, S_n).$$

For the vector $\mathbf{z} = (z_1, \dots, z_n)$ where $z_i \sim N(0, 1)$ for $1 \leq i \leq n$, each element in $\mathbf{S}$ is given by the recursive formula:

$$S_i = S_{i-1} \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t} z_i\right\}, \qquad S_0 = x.$$

Pricing the option numerically is basically a matter of approximating the discounted expectation by taking the ordinary mean of $m$ simulations of the option, where $m$ is chosen to be some suitably large number. Higher values of $m$ yield more precise approximations to the option price.

Each simulation of $S_t$ differs only through the vector $\mathbf{z}$, so for the $m$ simulations $\mathbf{S}^1, \mathbf{S}^2, \ldots, \mathbf{S}^m$ we express simulation $j$, $1 \leq j \leq m$, as a function of $\mathbf{z}$ and $x$:

$$\mathbf{S}^j = \mathbf{S}(\mathbf{z}^j, x).$$

Denoting the payoff function for an option by $\Phi(\cdot)$, (e.g for a European option with strike $K$, $\Phi(S_T) = (S_T - K)^+$ or alternatively $\Phi(\mathbf{S}^j) = \max\big(S_n - K, 0\big)\big)$, and adopting a new notation for the price of the option: $u(x)$, we have the Monte Carlo approximation to the option price given by:

$$u(x) = \mathbb{E}[\Phi(S_T)] \approx \frac{1}{m} \sum_{j=1}^{m} \Phi\big(\mathbf{S}(\mathbf{z}^j, x)\big) \tag{1.4}$$

(where we used $r = 0$ to avoid discounting the price, which we will assume is the case from now).

To calculate the Greeks, a careful choice of some small value $\varepsilon > 0$ is made (discussed in [12]), and by calculating $u(x+\varepsilon)$ and $u(x)$ using (1.4), we get the forward differencing approximation:

$$\Delta = \frac{\partial u(x)}{\partial x} \approx \frac{u(x+\varepsilon) - u(x)}{\varepsilon}.$$

An alternative is the centre differencing approach. This gives us a natural way to extend to the second derivative, in addition to improving the accuracy:

$$\Delta = \frac{\partial u(x)}{\partial x} \approx \frac{u(x+\varepsilon) - u(x-\varepsilon)}{2\varepsilon},$$

and the centre difference method for $\Gamma$ is:

$$\Gamma = \frac{\partial^2 u(x)}{\partial x^2} \approx \frac{u(x+\varepsilon) - 2u(x) + u(x-\varepsilon)}{\varepsilon^2}.$$

As they are estimates, the parameters are written with “hats”. From [12] the algorithms to numerically approximate $\Delta$ and $\Gamma$ are:

$$\widehat{\Delta} = \frac{1}{2m\varepsilon} \sum_{j=1}^{m} \Big[ \Phi\big(\mathbf{S}(\mathbf{z}^j, x+\varepsilon)\big) - \Phi\big(\mathbf{S}(\mathbf{z}^j, x-\varepsilon)\big) \Big]$$

$$\widehat{\Gamma} = \frac{1}{m\varepsilon^2} \sum_{j=1}^{m} \Big[ \Phi\big(\mathbf{S}(\mathbf{z}^j, x+\varepsilon)\big) - 2 \cdot \Phi\big(\mathbf{S}(\mathbf{z}^j, x)\big) + \Phi\big(\mathbf{S}(\mathbf{z}^j, x-\varepsilon)\big) \Big].$$

The other Greeks, $\rho$ and $\nu$, can be calculated using a similar approach.

## The Likelihood Ratio Method

For some option that only depends on the price model $S_t$ at the time $T$, e.g a European option, the payoff function will be on the form $\Phi(S_T)$, and when finding the option price when the risk-less interest rate is $r = 0$, the option value is $u(x) = \mathbb{E}[\Phi(S_T)]$.

When calculating the derivative of the option value, the general idea in the likelihood ratio method introduced by [6], is to transfer all the parameter dependencies from the payoff function to the density function. For some parameter $\theta$, the derivative can be found by the following set of equalities.

$$\begin{aligned}
\frac{\partial}{\partial \theta}\mathbb{E}[\Phi(S_T)] &= \frac{\partial}{\partial \theta}\int_{\mathbb{R}^m} \Phi(z) f_\theta(z) dz \\
&= \int_{\mathbb{R}^m} \Phi(z)\Big(\frac{\partial}{\partial \theta} f_\theta(z)\Big) dz \\
&\overset{(\star)}{=} \int_{\mathbb{R}^m} \Phi(z)\Big(\frac{\partial}{\partial \theta}\log\big[f_\theta(z)\big]\Big) f_\theta(z) dz \\
&= \mathbb{E}\Big[\Phi(S_T)\Big(\frac{\partial}{\partial \theta}\log f_\theta(S_T)\Big)\Big] \\
&= \mathbb{E}[\Phi(S_T)\pi],
\end{aligned} \tag{1.5}$$

where $\pi$ is called a *weight*,

$$\pi = \Big(\frac{\partial}{\partial \theta}\log f_\theta(S_T)\Big). \tag{1.6}$$

In the $(\star)$-transition, we used that:

$$\frac{\partial}{\partial \theta}\log\big[f_\theta(z)\big] = \frac{1}{f_\theta(z)} \cdot \frac{\partial}{\partial \theta} f_\theta(z)$$

and when multiplying both sides with $f_\theta(z)$:

$$\frac{\partial}{\partial \theta}\log\big[f_\theta(z)\big] f_\theta(z) = \frac{\partial}{\partial \theta} f_\theta(z).$$

Calculating the weighted option price by Monte Carlo simulation (as done in [12]) now becomes similar to calculating the option price as in (1.4):

$$\widehat{\Delta} = \frac{1}{m}\sum_{j=1}^{m}\Big[\Phi\big(\mathbf{S}(\mathbf{z}^j, x)\big)\cdot \pi\Big] = \frac{1}{m}\sum_{j=1}^{m}\Big[\Phi\big(\mathbf{S}(\mathbf{z}^j, x)\big)\cdot \frac{\partial}{\partial \theta}\log f_\theta(\mathbf{S}(\mathbf{z}^j, x))\Big]. \tag{1.7}$$

The differentiated weight could be calculated analytically before implementing the algorithm, or it could be calculated numerically by using the forward or centre differencing methods.

## Efficiency and Extensions

As discussed in [13], the forward difference method gives a poor convergence rate of $(1/n)^4$, which basically means for every decimal point precision that is desired, the simulations must be increased by a factor of $10^4$, which is extremely costly. However by using the centre difference method the convergence rate is improved to $(1/n)^3$ and by using the variance control technique of common variables in addition, it is possible to attain a convergence rate of $(1/n)^2$, which is the best possible case for Monte Carlo simulations.

The main problem with the finite difference method is its inability to cope with discontinuous payoff functions, most notably the digital option e.g with payoff function: $\Phi(S_T) = \mathbf{1}_{S_T>K}$ for some value $K$. This problem is also present in second order derivatives of continuous payoff functions, such as when calculating $\Gamma$ for a standard European call option.

The likelihood ratio method provides a $(1/n)^2$ convergence rate and does not depend on whether the payoff function $\Phi(S_T)$ is discontinuous or not. This method finds a way of calculating the derivative of the option value that does not involve differentiating the payoff function, which is the primary advantage.

The drawback of the likelihood ratio method is that the density function $f_\theta(\cdot)$ must be known, which is not always the case, and preferably that the density function is analytically differentiable, to avoid costly numerical approximations.

In 1999 Fournié et al. introduced another approach to calculating the Greeks by using Malliavin calculus to derive weights in a similar fashion to the likelihood ratio method. In [13] they showed how it is possible to derive a weight $\pi$:

$$\frac{\partial}{\partial\theta}\mathbb{E}[\Phi(S_T)] = \mathbb{E}[\Phi(S_T)\pi],$$

without needing to know the density function $f_\theta(\cdot)$. This method gives a $(1/n)^2$ convergence rate and is possible to apply even when the payoff function is discontinuous, eliminating the weaknesses of both the finite difference method and the likelihood ratio method.

One disadvantage of the new method is the rather high level of analytical calculations required, which depend on Malliavin calculus, an extension of the traditional Ito stochastic calculus. Deriving the weights for $\Delta$, $\Gamma$, $\rho$ and $\nu$ in this way will the main topic of Chapter 2.

In the cases where the finite difference method performs well, there isn't really any improvement when applying the new Malliavin calculus method. In fact, the finite difference method would be easier to implement as it does not rely on any advanced calculations beforehand. In short, the finite difference method is still preferable in certain situations.

## 1.3 Lévy Processes

The initial attempts at providing a mathematical model for the evolution of stock prices used Brownian motion to simulate the randomness in the market. The very first model as stated in [17] was introduced by Bachelier (1900):

$$S_t = S_0(1 + \sigma W_t). \tag{1.8}$$

There is also the well known geometric Brownian motion, introduced by Samuelson (1965) with dynamics given in (1.1) and solution (derived in (LA.2)) given by:

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t - \sigma W_t\right\}. \tag{1.9}$$

The problem with (1.8) is that the model permits negative stock prices, which of course is impossible. Geometric Brownian motion is always positive, but when compared to empirical data, it becomes apparent (1.9) does not give a realistic representation of how real world stock prices behave.

In various situations, e.g in stock market crashes or following disastrous news, stock prices jump: the value deemed by the market changes in an instant, and the stock price has a discontinuity. The model in (1.9) is not able to properly account for jumps of a certain magnitude since it is a continuous process. Even though it is theoretically possible for (1.9) to closely imitate jumps by e.g a very rapid decline, these movements are so unlikely they do not affect the model. As large, downward jumps occasionally happen, in practice this means that decisions based on continuous models may not have properly taken into account the potential downward risk, and option prices based on the same models may have been miscalculated.

To construct accurate models the randomness cannot be modelled by Brownian motion alone. One possible approach is to model the jumps by Lévy processes; a class of stochastic processes that includes Brownian motion as a special case. The only continuous Lévy process is Brownian motion; all the others are driven by jumps. The following definition of Lévy processes as given in [9].

**Definition 1.1** (Lévy Processes)
*A càdlàg stochastic process* $\{X_t \mid t \geq 0\}$ *on the probability space* $(\Omega, \mathcal{F}, P)$ *with values in* $\mathbb{R}$ *such that* $X_0 = 0$ *is called a* Lévy process *if it satisfies the following conditions:*

1. *Independent increments: for every increasing sequence of times* $t_0, \ldots, t_n$ *the random variables* $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ *are independent.*

2. *Stationary increments: the law of* $X_{t+h} - X_t$ *does not depend on* $t$.

3. *Stochastic continuity:* $\forall \varepsilon > 0, \ \lim_{h \to 0} P(|X_{t+h} - X_t| \geq \varepsilon) = 0.$

The càdlàg property, also called RCLL for "right continuous with left limits", describes the behaviour of the process at the jumps. If the Lévy process $X_t$ jumps at time $t$, we denote the size of the jump as

$$\Delta X_t = X_{t+} - X_{t-},$$

where $t+$ and $t-$ are the times directly after and before the jump, respectively. If we assume the càdlàg property, we have $X_t = X_{t+}$, or more informally: at time $t$ the process jumps first and then settles at a point.

Actually it is common to define Lévy processes without having the càdlàg property, but then it is possible to prove that the process has a unique modification that is càdlàg. Instead of assuming that we use the càdlàg modification, we can simply include it in the definition without loss of generality. (We say that $X_t$ is a modification of $Y_t$ if $P(X_t = Y_t) = 1$ for all $t \geq 0$).

An increment is the growth (or decline) of the process over a time interval. The independent increments property states that disjoint increments are independent random variables, which means the change in the process is independent of the previous behaviour. When the distribution of the increment only depends on the length of the interval throughout the process, it is said to have stationary increments, which enforces a loose type of uniform behaviour on the process. The third property, stochastic continuity, reflects the fact that we do not know when the jumps will come, since the jump times are random times.

Stock price models driven by Lévy processes are natural generalizations of (1.9). For a Lévy process $X_t$, the stock price is modelled as (like in e.g [16]):

$$S_t = S_0 \exp\{X_t\}, \tag{1.10}$$

where the special case $X_t = (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t$ is (1.9). Other than the traditional continuous case, there are in general two approaches to simulating financial models, as discussed in [9]. The first approach is the *jump diffusion* type, where the Lévy process has the form:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i. \tag{1.11}$$

The drift term $\gamma t$ and the Brownian motion term $\sigma W_t$ are the same as in the continuous case, but the third term (the sum) is the *compound Poisson process*. The random variable $N_t$ is the number given by a standard Poisson counting process with intensity $\lambda$, and the jump sizes $Y_i$ are identically distributed, independent random variables following some probability law, like for instance a Gaussian law $Y_i \sim N(0, a)$. The jump diffusion model can be simulated using more than one compound Poisson process or other Lévy processes. Two well known jump diffusion models are the Merton jump diffusion and Kou jump diffusion models.

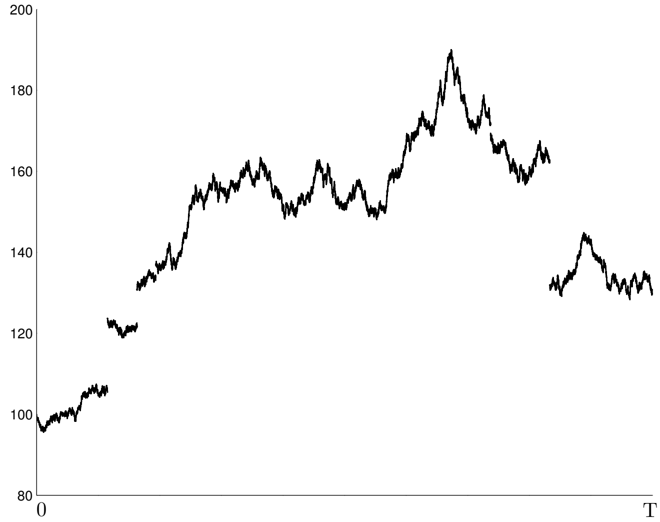


Figure 1.1: Jump diffusion.

A simulation of a jump diffusion model is depicted in Figure 1.1, where the trajectory is a financial model of the type (1.10) with $X_t$ as in (1.11), with $S_0 = 100$, $\gamma = 0$, $\sigma = 0.2$, $N_t$ a Poisson process with intensity $\lambda = 3$ and $Y_i \sim N(0, 0.1)$. The Brownian motion was simulated using `bmotion.m`, and the compound Poisson process using `compoisson.m`, both codes on page 56.

The second approach is using a process $X_t$ of the pure jump *infinite activity* type, which is a Lévy process that jumps infinitely often and which has been shown to be able to accurately describe properties of historical price processes. A representation of the form of an infinite activity model based on the Lévy-Ito decomposition is given in [9] as e.g:

$$X_t = \gamma t + \sum_{s \le t} \Delta X_s \mathbf{1}_{|\Delta X_s| \ge 1} + \lim_{\varepsilon \to 0} N_t^{\varepsilon},$$

where the "small jumps" are collected in the last term,

$$N_t^{\varepsilon} = \sum_{s \le t} \Delta X_s \mathbf{1}_{\varepsilon \le |\Delta X_s| < 1} - t \int_{\varepsilon \le |x| \le 1} x \nu(dx).$$

There is no Brownian motion term as the infinite activity models are flexible enough to capture nontrivial small time behaviour. A simulation of an exponential Lévy model with randomness modelled by the normal inverse Gaussian process is included in figure 1.2 on the following page, and is generated by the code included in `NIGP.m` on page 57. The parameters used are based on the ones given in [3]: $\delta = 0.0295$, $\alpha = 136.29$ and $\beta = -15.1977$.

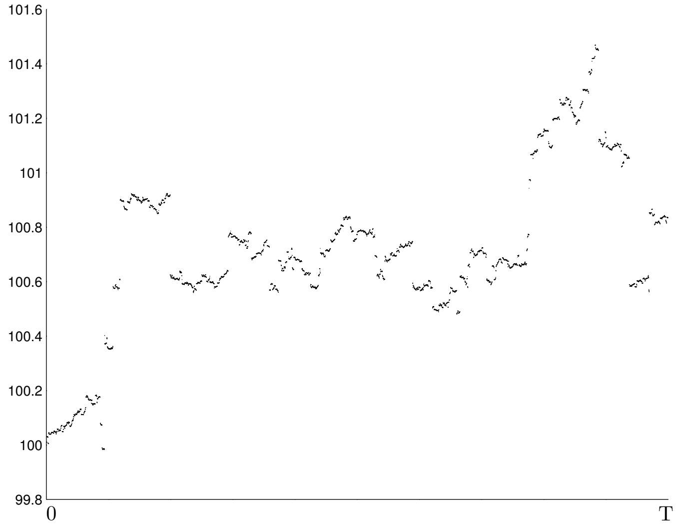


Figure 1.2: Infinite activity NIG process.

There are also pure jump Lévy models of finite activity, which would be like in the jump diffusion case without the Brownian motion, but as noted in [9], these models give a poor representation of the evolution of real world stocks and are of little interest.

Hedging financial positions remains important for discontinuous models, which includes the ability to calculate or numerically approximate the Greeks. The article ([13]) mentioned in section 1.2, that introduced Greeks calculated through Malliavin calculus, only derived weights for continuous models, but following the discussion of Lévy models above, the necessity of extending the results to discontinuous models becomes apparent.

# Chapter 2

# Malliavin Calculus and Expressions for the Greeks

The main goal of this chapter is to derive the central results from [13]. The results rely on Malliavin Calculus, so the first part of this chapter will be to give a summary of the theorems that will be needed as well as the definitions on which they depend. The presentation given here relies heavily on [11].

In section 2.2 there is a thorough discussion on [13], where the weights mentioned in Chapter 1, the Malliavin weights, are derived for some of the most important Greeks. Section 2.3 discusses some additional properties on the weights, as presented in [14].

## 2.1 Malliavin Calculus: Central Results

Throughout this section we denote the standard Brownian motion by $W_t$ for $t \in [0, T]$, and work with the complete probability space $(\Omega, \mathcal{F}, P)$ such that $W_0 = 0$ $P$-a.s. (The probability space is complete in the sense that it contains all subsets of $\Omega$ with $P$-outer measure zero).

The $\sigma$-algebra generated by Brownian motion $W_t$ is denoted by $\mathcal{F}_t$.

**Definition 2.1** (Iterated Ito integrals)
*For a symmetric, square integrable function $g(t_1, \ldots, t_n)$, we define the $n$-fold iterated Ito integral as:*

$$I_n(g) := \int_0^T \cdots \int_0^T g(t_1, \ldots, t_n) dW_{t_1} \ldots dW_{t_n}.$$

**Theorem 2.2** (The Wiener-Ito Chaos Expansion)
*Let $F$ be an $\mathcal{F}_T$-measurable random variable such that $\left(\mathbb{E}[F^2]\right)^{\frac{1}{2}} < \infty$. Then there exists a sequence of symmetric, square integrable functions $\{f_n\}_{n=0}^{\infty}$ on $[0,T]$ such that*

$$F = \sum_{n=0}^{\infty} I_n(f_n). \tag{2.1}$$

*Proof.*
Theorem 1.10 in [11]. ■

For a symmetric function $f_n = f_n(t_1, \ldots, t_n)$ we will sometimes be required to add an additional parameter to the function, so we get $f_n(t_1, \ldots, t_n, t) = f_n(\cdot, t) = f_{n,t}$. The extended function is no longer symmetric, so we define its symmetrization to be $\widetilde{f}_n = \widetilde{f}(t_1, \ldots, t_{n+1})$.

**Definition 2.3** (The Skorohod Integral)
*Let $u(t)$ be a measurable stochastic process such that $u(t)$ is $\mathcal{F}_T$-measurable for all $t \in [0,T]$ and $\mathbb{E}[u^2(t)] < \infty$, and assume its Wiener-Ito chaos expansion is*

$$u(t) = \sum_{n=0}^{\infty} I_n(f_n(\cdot, t)).$$

*Then we define the Skorohod integral of $u$ by:*

$$\delta(u) := \int_0^T u(t)\delta W_t := \sum_{n=0}^{\infty} I_{n+1}(\widetilde{f}_n),$$

*when this sum converges in $L^2(P)$, in which case we write $u \in Dom(\delta)$.*

A very useful property of the Skorohod integral is that it contains a class of Ito integrals when the integrand $u(t)$ is adapted with respect to the filtration $\mathcal{F}_t$ (i.e $u(t)$ is $\mathcal{F}_t$-measurable for all $t \in [0,T]$) as seen in the next theorem.

**Theorem 2.4**
*If $u(t)$, $t \in [0,T]$ is an adapted, measurable stochastic process such that $\|u\|^2_{L^2(P\times\lambda)} < \infty$ and $u(t)$ is Skorohod integrable: $u(t) \in Dom(\delta)$. Then the Skorohod integral coincides with the Ito integral:*

$$\int_0^T u(t)\delta W_t = \int_0^T u(t)dW_t.$$

*Proof.*
Theorem 2.9 in [11]. ■

**Definition 2.5**
*We define* $\mathbb{D}_{1,2} \subset L^2(P)$ *to be the set of Malliavin differentiable random variables. Let* $F \in L^2(P)$ *be* $\mathcal{F}_T$*-measurable with chaos expansion as given in (2.1). We say* $F \in \mathbb{D}_{1,2}$ *if*

$$\|F\|^2_{\mathbb{D}_{1,2}} := \sum_{n=1}^{\infty} nn! \|f_n\|^2_{L^2([0,T])^n} < \infty.$$

**Definition 2.6** (The Malliavin Derivative)
*If* $F \in \mathbb{D}_{1,2}$ *has a chaos expansion as in (2.1), we define the Malliavin derivative* $D_t F$ *of* $F$ *at time* $t$ *to be*

$$D_t F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)), \qquad t \in [0, T].$$

**Theorem 2.7** (The Chain Rule)
*We assume* $F \in \mathbb{D}_{1,2}$ *and that the function* $g$ *is differentiable with a bounded derivative. Then* $g(F) \in \mathbb{D}_{1,2}$*, and*

$$D_t g(F) = g'(F) D_t F.$$

*Proof.*
Theorem 3.5 in [11]. ∎

**Theorem 2.8** (The Duality Formula)
*Let* $F \in \mathbb{D}_{1,2}$ *be* $\mathcal{F}_T$*-measurable and let* $u(t)$ *be a Skorohod integrable stochastic process. Then*

$$\mathbb{E}\Big[F \int_0^T u(t) \delta W_t\Big] = \mathbb{E}\Big[\int_0^T u(t) D_t F dt\Big].$$

*Proof.*
Theorem 3.14 in [11]. ∎

**Theorem 2.9** (Integration by parts)
*Let* $u(t)$ *be a Skorohod integrable stochastic process and* $F \in \mathbb{D}_{1,2}$ *be such that* $Fu(t) \in Dom(\delta)$*. Then:*

$$\delta(Fu(t)) = F\delta(u(t)) - \int_0^T u(t) D_t F dt.$$

*Proof.*
Theorem 3.15 in [11]. ∎

**Theorem 2.10** (The Clark-Ocone Formula)
*Let $F \in \mathbb{D}_{1,2}$ be $\mathcal{F}_T$-measurable. Then*

$$F = \mathbb{E}[F] + \int_0^T \mathbb{E}[D_t F | \mathcal{F}_t] dW_t.$$

*Proof.*
Theorem 4.1 in [11]. ■

## 2.2 Malliavin Weights for the Greeks

Based on [13], with [5] as supporting reference, we will now derive the Malliavin weights for the Greeks. We will adopt the notation used in the first article. The first assumption we make is that the underlying financial asset is modelled by the process $\{X_t \mid t \in [0,T]\}$, and that this model satisfies the stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \qquad X_0 = x, \tag{2.2}$$

where $W_t$ is the standard one-dimensional Brownian motion and the initial value is some constant $x \in \mathbb{R}$. We assume that the functions $\mu(\cdot)$ and $\sigma(\cdot) > 0$ are continuously differentiable with bounded Lipschitz derivatives, in order to guarantee the existence of a strong solution, in which case $X_T \in \mathbb{D}_{1,2}$.

In [13] the payoff function depends on $m$ states of the underlying financial asset, but we make a slight simplification and only consider payoff functions depending on the terminal point $X_T$. We also restrict ourselves to the one dimensional case, so we avoid the multidimensional technicalities.

We denote the payoff function by $\Phi(\cdot)$ which can be regarded as a European option or a digital option. Following the notation in [13], we denote the value of the option (or even a contingent claim) as

$$u(x) = \mathbb{E}[\Phi(X_T) \mid X_0 = x] = \mathbb{E}^x[\Phi(X_T)]. \tag{2.3}$$

As done in [11], we will transfer the condition of $X_0 = x$ to the process, so we can write:

$$\mathbb{E}^x[\Phi(X_T)] = \mathbb{E}[\Phi(X_T^x)].$$

We assume from now on we have the following condition for the payoff function:

$$\|\Phi(X_T^x)\|^2_{L^2(P)} = \mathbb{E}[\Phi(X_T^x)^2] < \infty, \tag{2.4}$$

and in addition $\Phi(\cdot)$ is assumed to have a bounded derivative in order to allow the usage of the chain rule (T2.7).

We will require the *first variational process* $Y_t$, defined as $Y_t := \frac{\partial}{\partial x} X_t$, with dynamics given in [11] or found simply by differentiating (2.2) with respect to $x$,

$$dY_t = b'(X_t)Y_t dt + \sigma'(X_t)Y_t dW_t, \qquad Y_0 = 1. \tag{2.5}$$

Proceeding as in [13] we will derive the Malliavin weights for the Greeks introduced in Chapter 1, $\Delta$, $\rho$ and $\nu$. In addition the weight for $\Gamma$ is proved. In [13] the proofs are only sketched, but here they are given in full detail.

### 2.2.1 Delta

We define the set of square integrable functions $a$ whose integral over $[0, T]$ equals 1 as:

$$\mathcal{A} := \Big\{ a \in L^2([0,T]) \mid \int_0^T a(t)dt = 1 \Big\}, \tag{2.6}$$

where the typical choice will be $a(t) = \frac{1}{T}$.

For $\Delta$ we will require four additional supporting lemmas. The first lemma gives us conditions that allow us to change the order of the expectation and the derivative.

**Lemma 2.11**
*Suppose $F^\theta \in \mathbb{R}$ is a random variable that depends on some parameter $\theta \in \mathbb{R}$, and suppose for almost every $\omega \in \Omega$ that the mapping $\theta \mapsto F^\theta(\omega)$ is continuously differentiable in $[a, b]$ and that*

$$\mathbb{E}\Big[ \sup_{\theta \in [a,b]} \Big| \frac{\partial F^\theta}{\partial \theta} \Big| \Big] < \infty.$$

*Then the mapping $\theta \mapsto \mathbb{E}[F^\theta]$ is differentiable in $(a, b)$, and for $\theta \in (a, b)$ we can change the order of the derivative and the expectation:*

$$\frac{\partial}{\partial \theta}\mathbb{E}[F^\theta] = \mathbb{E}\Big[ \frac{\partial}{\partial \theta} F^\theta \Big].$$

*Proof.*
Lemma 4.1 in [4]. ■

In (LA.9) on page 54 we show that $F^\theta = \Phi(X_T^x)$ satisfies (L2.11) when $\theta = x$, and note that the other cases can be shown in a similar manner.

The following lemma allows us to assume a smoothness condition for the payoff function $\Phi$. We denote the price model $X_t$ by $X_t^\theta$ to signify the dependence on some parameter $\theta$.

**Lemma 2.12**
*Let $\theta \mapsto \pi^\theta$ be a process such that $\theta \mapsto \psi(\theta) := \|\pi^\theta\|_{L^2(P)}$ is locally bounded. Assume that:*

$$\frac{\partial}{\partial\theta}\mathbb{E}[\Phi(X_T^\theta)] = \mathbb{E}[\Phi(X_T^\theta)\pi^\theta]$$

*is valid for all $\Phi \in C_c^\infty(\mathbb{R})$ (infinitely differentiable with compact support). Then we can extend this equality to all $\Phi \in L^2(\mathbb{R})$.*

*Proof.*
Lemma 12.28 in [11] or Lemma 4.2 in [4]. ■

The next two lemmas provide some necessary equalities.

**Lemma 2.13**
*An alternative expression for the Malliavin derivative of $X_s$. $Y_t$ denotes the first variational process (2.5).*

$$D_s X_t = \frac{Y_t}{Y_s}\sigma(X_s)\mathbf{1}_{[0,t]}(s)$$

*Proof.*
Lemma 4.16 in [11]. ■

**Lemma 2.14**
*Let $a \in \mathcal{A}$ as in (2.6). Then*

$$Y_T = \int_0^T D_s X_T \frac{Y_s}{\sigma(X_s)} a(s) ds.$$

*Proof.*

$$Y_T = Y_T \cdot 1 = Y_T \int_0^T a(s)ds = \int_0^T Y_T a(s) ds. \tag{2.7}$$

By (L2.13) we have:

$$D_s X_T = \frac{Y_T}{Y_s}\sigma(X_s)\mathbf{1}_{[0,T]}(s)$$

and when solved for $Y_T$, and using $\mathbf{1}_{[0,T]}(s) = 1$, we get:

$$Y_T = D_s X_T \frac{Y_s}{\sigma(X_s)}. \tag{2.8}$$

Completing the proof using these two equations:

$$Y_T \overset{(2.7)}{=} \int_0^T Y_T a(s) ds \overset{(2.8)}{=} \int_0^T D_s X_T \frac{Y_s}{\sigma(X_s)} a(s) ds.$$

■

Finally we have the necessary setup to derive the Malliavin weight for $\Delta$.

**Proposition 2.15** (Malliavin weight for $\Delta$)
*For any $x \in \mathbb{R}$ and any $a \in \mathcal{A}$, we have:*

$$\frac{\partial}{\partial x} u(x) = \mathbb{E}^x\Big[\Phi(X_T) \int_0^T \frac{a(t)Y_t}{\sigma(X_t)} dW_t\Big],$$

*so the Malliavin weight for $\Delta$ is $\pi^\Delta = \int_0^T \frac{a(t)Y_t}{\sigma(X_t)} dW_t$.*

*Proof.*
We can prove this result using the following set of equalities. Assuming $\Phi$ is infinitely differentiable.

$$\begin{aligned}
\frac{\partial}{\partial x} u(x) &\overset{(2.3)}{=} \frac{\partial}{\partial x}\mathbb{E}^x[\Phi(X_T)] \\
&= \frac{\partial}{\partial x}\mathbb{E}[\Phi(X_T^x)] \\
&\overset{(L2.11)}{=} \mathbb{E}\Big[\frac{\partial}{\partial x}\Phi(X_T^x)\Big] \\
&= \mathbb{E}[\Phi'(X_T^x)Y_T] \\
&\overset{(L2.14)}{=} \mathbb{E}\Big[\Phi'(X_T^x) \int_0^T D_s X_T^x \frac{a(s)Y_s}{\sigma(X_s^x)} ds\Big] \\
&= \mathbb{E}\Big[\int_0^T \Phi'(X_T^x) D_s X_T^x \frac{a(s)Y_s}{\sigma(X_s^x)} ds\Big] \\
&\overset{(T2.7)}{=} \mathbb{E}\Big[\int_0^T D_s \Phi(X_T^x) \frac{a(s)Y_s}{\sigma(X_s^x)} ds\Big] \\
&\overset{(T2.8)}{=} \mathbb{E}\Big[\Phi(X_T^x) \int_0^T \frac{a(s)Y_s}{\sigma(X_s^x)} \delta W_s\Big] \\
&\overset{(T2.4)}{=} \mathbb{E}\Big[\Phi(X_T^x) \int_0^T \frac{a(s)Y_s}{\sigma(X_s^x)} dW_s\Big] \\
&= \mathbb{E}^x\Big[\Phi(X_T) \int_0^T \frac{a(s)Y_s}{\sigma(X_s)} dW_s\Big] \\
&= \mathbb{E}^x\big[\Phi(X_T)\pi^\Delta\big].
\end{aligned}$$

By Lemma (L2.12) this result also applies to all $\Phi \in L^2(\mathbb{R})$. ■

### 2.2.2 Gamma

In addition to the assumptions for $\Delta$, we assume that $\mu(\cdot)$ and $\sigma(\cdot)$ have bounded second order derivatives. A process that will be needed is the *second variation process*, $U_t := \frac{\partial}{\partial x} Y_t = \frac{\partial^2}{\partial x^2} X_t$, with dynamics given by:

$$dU_t = \Big(b'(X_t)U_t + b''(X_t)Y_t^2\Big)dt + \Big(\sigma'(X_t)U_t + \sigma''(X_t)Y_t^2\Big)dW_t, \qquad U_0 = 0,$$

which we get by differentiating (2.5) by $x$ or consulting [8].

The next two lemmas will make the proof of the main Proposition a lot shorter.

**Lemma 2.16**

$$\frac{\partial}{\partial x}\pi^{\Delta} = \int_0^T a(s)\frac{U_s\sigma(X_s) - Y_s\sigma'(X_s)}{\sigma^2(X_s)}dW_s =: G_s$$

*Proof.*

$$\begin{aligned}
\frac{\partial}{\partial x}\pi^{\Delta} &\overset{(P2.15)}{=} \frac{\partial}{\partial x}\int_0^T \frac{a(s)Y_s}{\sigma(X_s)}dW_s \\
&= \int_0^T a(s)\frac{\partial}{\partial x}\frac{Y_s}{\sigma(X_s)}dW_s \\
&= \int_0^T a(s)\frac{U_s\sigma(X_s) - Y_s\sigma'(X_s)}{\sigma^2(X_s)}dW_s
\end{aligned}$$

■

**Lemma 2.17**

$$\mathbb{E}[\pi^{\Delta}\Phi'(X_T)Y_T] = \mathbb{E}\Big[\Phi(X_T)\underbrace{\Big((\pi^{\Delta})^2 - \int_0^T \Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big)^2 ds - H_s\Big)}_{=:F_s}\Big]$$

*Proof.*

$$\begin{aligned}
\mathbb{E}[\pi^{\Delta}\Phi'(X_T)Y_T] &\overset{(L2.14)}{=} \mathbb{E}\Big[\pi^{\Delta}\Phi'(X_T)\int_0^T D_sX_T\frac{a(s)Y_s}{\sigma(X_s)}ds\Big] \\
&\overset{(T2.7)}{=} \mathbb{E}\Big[\int_0^T \pi^{\Delta}D_s\Phi(X_T)\frac{a(s)Y_s}{\sigma(X_s)}ds\Big] \\
&\overset{(T2.8)}{=} \mathbb{E}\Big[\Phi(X_T)\delta\Big(\pi^{\Delta}\frac{a(s)Y_s}{\sigma(X_s)}\Big)\Big] \\
&\overset{(T2.9)}{=} \mathbb{E}\Big[\Phi(X_T)\Big(\pi^{\Delta}\delta\Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big) - \int_0^T \frac{a(s)Y_s}{\sigma(X_s)}D_s\pi^{\Delta}ds\Big)\Big] \\
&= \mathbb{E}\Big[\Phi(X_T)\Big((\pi^{\Delta})^2 - \int_0^T \Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big)^2 ds - H_s\Big)\Big]
\end{aligned}$$

The last step follows since

$$\delta\Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big) \overset{(T2.4)}{=} \int_0^T \frac{a(s)Y_s}{\sigma(X_s)}dW_s = \pi^\Delta,$$

and since $u(s)$ is adapted, we can apply Corollary 3.19 in [11]:

$$D_s\pi^\Delta = D_s\Big(\int_0^T \frac{a(r)Y_r}{\sigma(X_r)}dW_r\Big) = \frac{a(s)Y_s}{\sigma(X_s)} + \int_s^T D_s\frac{a(r)Y_r}{\sigma(X_r)}dW_r \implies$$

$$\begin{aligned}\int_0^T \frac{a(s)Y_s}{\sigma(X_s)}D_s\pi^\Delta ds &= \int_0^T \Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big)^2 ds + \underbrace{\int_0^T \frac{a(s)Y_s}{\sigma(X_s)}\Big(\int_s^T D_s\frac{a(r)Y_r}{\sigma(X_r)}dW_r\Big)ds}_{=:H_s} \\ &= \int_0^T \Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big)^2 ds + H_s. \qquad (2.9)\end{aligned}$$

■

**Proposition 2.18** (Malliavin weight for Γ)
*For any $x \in \mathbb{R}$ and any $a \in \mathcal{A}$:*

$$\pi^\Gamma = \underbrace{(\pi^\Delta)^2 - \int_0^T \Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big)^2 ds - H_s}_{F_s} + \underbrace{\int_0^T a(s)\frac{U_s\sigma(X_s) - Y_s\sigma'(X_s)}{\sigma^2(X_s)}dW_s}_{G_s}$$

*Proof.*

$$\begin{aligned}\frac{\partial^2 u(x)}{\partial x^2} &= \frac{\partial^2}{\partial x^2}\mathbb{E}[\Phi(X_T^x)] \\ &\overset{(P2.15)}{=} \frac{\partial}{\partial x}\mathbb{E}[\Phi(X_T^x)\pi^\Delta] \\ &\overset{(L2.11)}{=} \mathbb{E}\Big[\pi^\Delta\frac{\partial}{\partial x}\Phi(X_T^x) + \Phi(X_T^x)\frac{\partial}{\partial x}\pi^\Delta\Big] \\ &= \mathbb{E}\Big[\pi^\Delta\Phi'(X_T^x)Y_T + \Phi(X_T^x)\frac{\partial}{\partial x}\pi^\Delta\Big] \\ &\overset{(L2.17)}{=} \mathbb{E}\Big[\Phi(X_T^x)F_s + \Phi(X_T^x)\frac{\partial}{\partial x}\pi^\Delta\Big] \\ &\overset{(L2.16)}{=} \mathbb{E}[\Phi(X_T^x)F_s + \Phi(X_T^x)G_s] \\ &= \mathbb{E}[\Phi(X_T^x)(F_s + G_s)]\end{aligned}$$

■

### 2.2.3 Rho

The two previous Malliavin weights were derived in a similar manner using Malliavin calculus, but $\pi^\rho$ requires a different approach. The weight is found by calculating the Gateaux derivative, a generalization of the partial derivative to Banach spaces (complete normed vector spaces), which is done in the drift direction through a perturbed process. The perturbed stochastic differential equation is the original equation (2.2) with a small length added in the drift direction. By using Girsanov's theorem the perturbed process is reduced to the original stochastic differential equation where we can derive the weight.

The proof of this result as presented in [13] contains two errors, but the general approach is correct and the proof can be corrected with two small adjustments. Because the proof given here is a correction and is largely based on classical stochastic calculus, the level of detail will be somewhat higher.

For the main proposition there will be need for the following two results from standard measure theory.

**Theorem 2.19** (Jensen's inequality)
*Assume $f(x) : \mathbb{R} \mapsto \mathbb{R}$ is a convex function (e.g $f(x) = |x|$) , $X \in L^1(P)$. Then*

$$f\big(\mathbb{E}[X]\big) \leq \mathbb{E}[f(X)].$$

*Proof.*
Theorem 12.14 in [18]. ■

**Theorem 2.20** (The Cauchy-Schwarz inequality)
*Assume $X, Y \in L^2(P)$. Then $X \cdot Y \in L^1(P)$ and*

$$\|XY\|_{L^1(P)} = \mathbb{E}[|XY|] \leq \big(\mathbb{E}[|X|^2]\mathbb{E}[|Y|^2]\big)^{\frac{1}{2}} = \|X\|_{L^2(P)}\|Y\|_{L^2(P)}.$$

*Proof.*
Corollary 12.3 in [18]. ■

For some variable $\varepsilon > 0$ and some bounded function $\gamma : [0,T] \times \mathbb{R} \mapsto \mathbb{R}$, the *perturbed process* $X_t^\varepsilon$ is defined by its dynamics:

$$dX_t^\varepsilon = \big[b(X_t^\varepsilon) + \varepsilon\gamma(X_t^\varepsilon)\big]dt + \sigma(X_t^\varepsilon)dW_t, \qquad X_0^\varepsilon = x, \tag{2.10}$$

where we note that $\varepsilon = 0$ returns us to $X_t$ as in (2.2). Associated with the perturbed process (2.10) is the perturbed option value:

$$u^\varepsilon(x) = \mathbb{E}^x[\Phi(X_T^\varepsilon)]. \tag{2.11}$$

We define the random variable:

$$Z_T^\varepsilon = \exp\Big\{ -\varepsilon \int_0^T \frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)} dW_t - \frac{\varepsilon^2}{2} \int_0^T \Big(\frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)}\Big)^2 dt \Big\}. \tag{2.12}$$

Since $\gamma(X_t^\varepsilon)$ is assumed to be bounded on $t \in [0, T]$, and $\sigma(X_t^\varepsilon) \geq \alpha > 0$ for some $\alpha \in \mathbb{R}$ and all $t \in [0, T]$, it follows that $\frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)}$ is finite on $t \in [0, T]$, so the Novikov condition (from e.g [21]) holds:

$$\mathbb{E}\Big[ \exp \Big\{ \frac{1}{2} \int_0^T \Big| \frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)} \Big|^2 dt \Big\} \Big] < \infty,$$

which is a sufficient condition for (2.12) to be a martingale. By the martingale property we can find the expectation of $Z_T^\varepsilon$:

$$\mathbb{E}[Z_T^\varepsilon] = \mathbb{E}[Z_T^\varepsilon \mid \mathcal{F}_0] = \mathbb{E}[Z_0^\varepsilon] = \mathbb{E}[e^0] = e^0 = 1. \tag{2.13}$$

Now we can move on to the main result of this subsection.

**Proposition 2.21** (Malliavin weight for $\rho$)
*The function $\varepsilon \mapsto u^\varepsilon(x)$ is differentiable in $\varepsilon = 0$ for any $x \in \mathbb{R}$, and we have:*

$$\frac{\partial}{\partial \varepsilon} u^\varepsilon(x) \Big|_{\varepsilon=0} = \mathbb{E}^x \Big[ \Phi(X_T) \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t \Big],$$

*so* $\pi^\rho = \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t$.

*Proof.*
We begin by defining the new probability measure $Q^\varepsilon$ by

$$dQ^\varepsilon := Z_T^\varepsilon dP,$$

where $Z_T^\varepsilon$ (as in (2.12)) is the Radon-Nikodym derivative of $Q^\varepsilon$ with respect to $P$. By properties of the Radon-Nikodym derivative, $Q^\varepsilon$ is absolutely continuous with respect to $P$ (i.e for any set $H$, $P(H) = 0 \Rightarrow Q^\varepsilon(H) = 0$), which we denote as $Q^\varepsilon \ll P$.

By (2.13) we have $\mathbb{E}[Z_T^\varepsilon] = 1 > 0$ a.s, so by [21], $Q^\varepsilon \gg P$, thus they are equivalent probability measures, a relationship denoted as $P \sim Q^\varepsilon$.

By Girsanov's Theorem, we can define the Wiener process with regards to $Q^\varepsilon$ as:

$$W_t^\varepsilon := W_t + \varepsilon \int_0^t \frac{\gamma(X_s^\varepsilon)}{\sigma(X_s^\varepsilon)} ds \implies dW_t = dW_t^\varepsilon - \varepsilon \frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)} dt. \tag{2.14}$$

Applying Girsanov's theorem to the perturbed process (2.10) under the probability measure $Q^\varepsilon$:

$$
\begin{aligned}
dX_t^\varepsilon &= \big[b(X_t^\varepsilon) + \varepsilon\gamma(X_t^\varepsilon)\big]dt + \sigma(X_t^\varepsilon)dW_t \\
&\overset{(2.14)}{=} \big[b(X_t^\varepsilon) + \varepsilon\gamma(X_t^\varepsilon)\big]dt + \sigma(X_t^\varepsilon)\Big(dW_t^\varepsilon - \varepsilon\frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)}dt\Big) \\
&= \big[b(X_t^\varepsilon) + \varepsilon\gamma(X_t^\varepsilon) - \varepsilon\gamma(X_t^\varepsilon)\big]dt + \sigma(X_t^\varepsilon)dW_t^\varepsilon \\
&= b(X_t^\varepsilon)dt + \sigma(X_t^\varepsilon)dW_t^\varepsilon.
\end{aligned}
$$

Since $X_0^\varepsilon = x$, we see that $X_t^\varepsilon$ follows the same stochastic differential equation under $Q^\varepsilon$ as the original process $X_t$ (2.2) does under $P$.

Since the probability measures are equivalent: $Q^\varepsilon \sim P$, it is possible to find the inverse of the Radon-Nikodym derivative. By a result in e.g [7]:

$$
\widetilde{Z}_T^\varepsilon := \frac{dP}{dQ^\varepsilon} = \Big(\frac{dQ^\varepsilon}{dP}\Big)^{-1} = (Z_T^\varepsilon)^{-1}.
$$

This is simply the inverse of $Z_T^\varepsilon$ (and in [13] the first term is erroneously negative):

$$
\widetilde{Z}_T^\varepsilon = \exp\Big\{\varepsilon\int_0^T \frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)}dW_t + \frac{\varepsilon^2}{2}\int_0^T \Big(\frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)}\Big)^2 dt\Big\},
$$

and after inserting the Brownian motion under $Q^\varepsilon$ as given in (2.14):

$$
\widetilde{Z}_T^\varepsilon = \exp\Big\{\varepsilon\int_0^T \frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)}dW_t^\varepsilon - \frac{\varepsilon^2}{2}\int_0^T \Big(\frac{\gamma(X_t^\varepsilon)}{\sigma(X_t^\varepsilon)}\Big)^2 dt\Big\}.
$$

Under the new probability measure we get a new expression for the perturbed option value (2.11):

$$
\begin{aligned}
u^\varepsilon(x) = \mathbb{E}^x[\Phi(X_T^\varepsilon)] &= \int_\Omega \Phi(X_T^\varepsilon(\omega))dP(\omega) \\
&= \int_\Omega \Phi(X_T^\varepsilon(\omega))\widetilde{Z}_T^\varepsilon dQ^\varepsilon(\omega) \quad = \quad \mathbb{E}_{Q^\varepsilon}^x[\Phi(X_T^\varepsilon)\widetilde{Z}_T^\varepsilon].
\end{aligned}
$$

Since the distribution of $(X_t^\varepsilon, W_t^\varepsilon)$ under $Q^\varepsilon$ coincides with $(X_t, W_t)$ under $P$, we can rewrite the perturbed option value as done in [13]:

$$
u^\varepsilon(x) \;=\; \mathbb{E}_{Q^\varepsilon}^x[\Phi(X_T^\varepsilon)\widetilde{Z}_T^\varepsilon] \;=\; \mathbb{E}^x[\Phi(X_T)\widehat{Z}_T^\varepsilon], \tag{2.15}
$$

where $\widehat{Z}_T^\varepsilon$ has the same form as $\widetilde{Z}_T^\varepsilon$ with $X_t$ and $W_t$ instead of $X_t^\varepsilon$ and $W_t^\varepsilon$:

$$
\widehat{Z}_T^\varepsilon = \exp\Big\{\varepsilon\int_0^T \frac{\gamma(X_t)}{\sigma(X_t)}dW_t - \frac{\varepsilon^2}{2}\int_0^T \Big(\frac{\gamma(X_t)}{\sigma(X_t)}\Big)^2 dt\Big\}.
$$

By Lemma A.3 on page 51, $\widehat{Z}_T^\varepsilon$ has the following integral form:

$$\widehat{Z}_T^\varepsilon = 1 + \varepsilon \int_0^T \widehat{Z}_t^\varepsilon \frac{\gamma(X_t)}{\sigma(X_t)} dW_t \implies \frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} = \int_0^T \widehat{Z}_t^\varepsilon \frac{\gamma(X_t)}{\sigma(X_t)} dW_t.$$

When taking the limit $\varepsilon \to 0$, both terms in the exponential tend to 0, so $\widehat{Z}_t^\varepsilon \to 1$:

$$\lim_{\varepsilon \to 0} \frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} = \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t. \tag{2.16}$$

(In [13] the original $Z_T^\varepsilon$ is used instead of the new $\widehat{Z}_T^\varepsilon$ which is the second mistake). Now we can finalize our proof with the following set of inequalities.

$$\begin{aligned}
& \left| \frac{1}{\varepsilon}\big(u^\varepsilon(x) - u(x)\big) - \mathbb{E}^x\left[\Phi(X_T) \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right] \right| \\
\overset{(2.15)}{=} \quad & \left| \frac{1}{\varepsilon}\Big(\mathbb{E}^x\big[\widehat{Z}_T^\varepsilon \Phi(X_T)\big] - \mathbb{E}^x[\Phi(X_T)]\Big) - \mathbb{E}^x\left[\Phi(X_T) \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right] \right| \\
= \quad & \left| \mathbb{E}^x\left[\frac{\widehat{Z}_T^\varepsilon \Phi(X_T) - \Phi(X_T)}{\varepsilon} - \Phi(X_T) \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right] \right| \\
= \quad & \left| \mathbb{E}^x\left[\Phi(X_T)\left(\frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} - \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right)\right] \right| \\
\overset{(T2.19)}{\leq} \quad & \mathbb{E}^x\left[\left|\Phi(X_T)\left(\frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} - \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right)\right|\right] \\
= \quad & \left\|\Phi(X_T)\left(\frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} - \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right)\right\|_{L^1(P)} \\
\overset{(T2.20)}{\leq} \quad & \big\|\Phi(X_T)\big\|_{L^2(P)} \left\|\frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} - \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right\|_{L^2(P)} \\
\overset{(2.4)}{\leq} \quad & K \cdot \left\|\frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} - \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right\|_{L^2(P)} \\
= \quad & K \cdot \mathbb{E}^x\left[\left(\frac{\widehat{Z}_T^\varepsilon - 1}{\varepsilon} - \int_0^T \frac{\gamma(X_t)}{\sigma(X_t)} dW_t\right)^2\right]^{\frac{1}{2}} \to 0 \text{ as } \varepsilon \to 0 \text{ by (2.16)}.
\end{aligned}$$

$K$ is some finite value such that $\|\Phi(X_T)\|_{L^2(P)} \leq K$ which we know exists by equation (2.4). As $\varepsilon \to 0$, the final term tends to 0, completing the proof. ∎

### 2.2.4 Vega

Proceeding in a similar way as for Rho, we define the perturbed process:

$$dX_t^\varepsilon = b(X_t^\varepsilon) + \big[\sigma(X_t^\varepsilon) + \varepsilon\widetilde{\sigma}(X_t^\varepsilon)\big]dW_t, \qquad X_0^\varepsilon = x,$$

where the we assume $\widetilde{\sigma}(X_t^\varepsilon) > 0$ and $\sigma(X_t^\varepsilon) + \varepsilon\widetilde{\sigma}(X_t^\varepsilon) > 0$ for all $t \in [0, T]$. As for the perturbed process used for Rho, when $\varepsilon = 0$ the stochastic differential equation becomes the same as in (2.2). Defining $Y_t^\varepsilon := \frac{\partial}{\partial x} X_t^\varepsilon$ yields the first variation process with respect to $x$, which is driven by the following dynamics:

$$dY_t^\varepsilon = b'(X_t^\varepsilon)Y_t^\varepsilon + \big[\sigma'(X_t^\varepsilon) + \varepsilon\widetilde{\sigma}'(X_t^\varepsilon)\big]Y_t^\varepsilon dW_t, \qquad Y_0^\varepsilon = 1.$$

And by defining $Z_t^\varepsilon = \frac{\partial}{\partial \varepsilon} X_t^\varepsilon$, the first variation process with respect to $\varepsilon$, the resulting stochastic differential equation is driven by:

$$dZ_t^\varepsilon = b'(X_t^\varepsilon)Z_t^\varepsilon + \widetilde{\sigma}(X_t^\varepsilon)dW_t + \big[\sigma'(X_t^\varepsilon) + \varepsilon\widetilde{\sigma}'(X_t^\varepsilon)\big]Z_t^\varepsilon dW_t \qquad Z_0^\varepsilon = 0.$$

When $\varepsilon = 0$ these three processes will be denoted as $X_t$, $Y_t$ and $Z_t$, respectively.

Based on these processes, we define $\beta(t) := \frac{Z_t}{Y_t}$ for $t \in [0, T]$, and note that we get the equality:

$$Z_T = \beta(T)Y_T \tag{2.17}$$

as well as $\beta(0) = \frac{Z_0}{Y_0} = 0$. Having the necessary setup, we can find the weight.

**Proposition 2.22** (Malliavin weight for $\nu$)
*For any $a \in \mathcal{A}$ and using $\widetilde{\beta}_a(T) := (\beta(T) - \beta(0))a(t) = \beta(T)a(t)$:*

$$\frac{\partial}{\partial \varepsilon} u^\varepsilon(x)\Big|_{\varepsilon=0} = \mathbb{E}^x\Big[\Phi(X_T)\delta\Big(\frac{Y_t}{\sigma(X_t)}\widetilde{\beta}_a(T)\Big)\Big].$$

*Proof.*

$$\begin{aligned}
\frac{\partial}{\partial \varepsilon} u^\varepsilon(x)\Big|_{\varepsilon=0} &= \frac{\partial}{\partial \varepsilon}\mathbb{E}^x[\Phi(X_T^\varepsilon)]\big|_{\varepsilon=0} \\
&\overset{(L2.11)}{=} \mathbb{E}^x[\Phi'(X_T^\varepsilon)Z_T^\varepsilon]\big|_{\varepsilon=0} \\
&= \mathbb{E}^x[\Phi'(X_T)Z_T] \\
&\overset{(2.17)}{=} \mathbb{E}^x[\Phi'(X_T)\beta(T)Y_T] \\
&\overset{(L2.14)}{=} \mathbb{E}^x[\int_0^T \Phi'(X_T)DtX_T\beta(T)a(t)\frac{Y_t}{\sigma(X_t)}dt] \\
&\overset{(T2.7)}{=} \mathbb{E}^x[\int_0^T D_t\Phi(X_T)\widetilde{\beta}_a(T)\frac{Y_t}{\sigma(X_t)}dt] \\
&\overset{(T2.8)}{=} \mathbb{E}^x\Big[\Phi(X_T)\delta\Big(\widetilde{\beta}_a(T)\frac{Y_t}{\sigma(X_t)}\Big)\Big].
\end{aligned}$$

The last step requires a confirmation that $\widetilde{\beta}_a(T)\frac{Y_t}{\sigma(X_t)}$ is Skorohod integrable, but we refer to [13] (which in turn refers to other sources) for the details. We can't apply $(T2.4)$, since the process is not adapted due to the dependence on $T$ in $\widetilde{\beta}_a(T)$. ■

## Example

As an application of the propositions, we will derive the Malliavin Weights $\pi^{\Delta}_{BS}$ and $\pi^{\Gamma}_{BS}$ in the Black-Scholes framework.

From the results above, we have the general Malliavin weights:

$$\begin{aligned}
\pi^{\Delta} &= \int_0^T \frac{a(t)Y_t}{\sigma(X_t)} dW_t \\
\pi^{\Gamma} &= (\pi^{\Delta})^2 + \int_0^T a(s)\frac{U_s\sigma(X_s) - Y_s\sigma'(X_s)}{\sigma^2(X_s)} dW_s - \int_0^T \Big(\frac{a(s)Y_s}{\sigma(X_s)}\Big)^2 ds - H_s
\end{aligned}$$

Under the risk neutral probability measure, the dynamics for the geometric Brownian motion is:

$$dS_t = r(t)S_t dt + \sigma S_t dW_t, \qquad S_0 = x,$$

for some interest rate model $r(t)$ and where we note that $\sigma(X_t) = \sigma S_t$. The solution is:

$$S_t = x\exp\Big\{\Big(r(t) - \frac{1}{2}\sigma^2\Big)t - \sigma W_t\Big\}.$$

Differentiating this with respect to $x$ gives the first variation process:

$$Y_t = \exp\Big\{\Big(r(t) - \frac{1}{2}\sigma^2\Big)t - \sigma W_t\Big\},$$

and we note the relation $Y_t = \frac{1}{x}S_t$. There is no dependence on $x$ in the first variation process, so $U_t = \frac{\partial}{\partial x}Y_t = 0$. We will use $a(t) = \frac{1}{T}$ in all the following calculations.

$$\begin{aligned}
\pi^{\Delta}_{BS} &= \int_0^T \frac{a(t)Y_t}{\sigma(X_t)} dW_t \\
&= \int_0^T \frac{1}{T}\frac{Y_t}{\sigma S_t} dW_t \\
&= \int_0^T \frac{1}{x\sigma T}\frac{\cancel{S_t}}{\cancel{S_t}} dW_t \qquad (2.18)\\
&= \frac{1}{x\sigma T}\int_0^T dW_t \\
&= \frac{W_T}{x\sigma T}
\end{aligned}$$

For $\pi^{\Gamma}_{BS}$ each term is considered separately. First term.

$$(\pi^{\Delta}_{BS})^2 \overset{(2.18)}{=} \Big(\frac{W_T}{x\sigma T}\Big)^2 = \frac{W_T^2}{x^2\sigma^2 T^2}. \tag{2.19}$$

The second term, using that $\sigma'(X_t) = \frac{\partial}{\partial x}\sigma S_t = \sigma Y_t$.

$$\begin{aligned} \int_0^T a(t)\frac{U_t\sigma(X_t) - Y_t\sigma'(X_t)}{\sigma^2(X_t)}dW_t &= \frac{1}{T}\int_0^T \frac{0-\sigma Y_t^2}{\sigma^2 S_t^2}dW_t \\ &= -\frac{1}{x^2\sigma T}\int_0^T \frac{S_t^2}{S_t^2}dW_t \\ &= -\frac{1}{x^2\sigma T}\int_0^T dW_t \\ &= -\frac{W_T}{x^2\sigma T} \end{aligned} \tag{2.20}$$

In passing we note that we could also write this term as $-\frac{1}{x}\pi^{\Delta}_{BS}$.

Third term.

$$\begin{aligned} -\int_0^T \Big(\frac{a(t)Y_t}{\sigma(X_t)}\Big)^2 dt &= -\int_0^T \Big(\frac{1}{T}\frac{Y_t}{\sigma S_t}\Big)^2 ds \\ &= -\frac{1}{x^2\sigma^2T^2}\int_0^T \frac{S_t^2}{S_t^2}ds \\ &= -\frac{1}{x^2\sigma^2T^2}\int_0^T ds \\ &= -\frac{1}{x^2\sigma^2 T} \end{aligned} \tag{2.21}$$

The fourth term, $H_s$, as defined in (2.9) on page 21, becomes 0:

$$D_s\Big(\frac{a(r)Y_r}{\sigma(X_r)}\Big) = D_s\Big(\frac{S_r}{x\sigma T S_r}\Big) = D_s\Big(\frac{1}{x\sigma T}\Big) = 0 \implies H_s = 0.$$

Finally we can combine (2.19), (2.20) and (2.21).

$$\begin{aligned} \pi^{\Gamma}_{BS} &= (\pi^{\Delta}_{BS})^2 + \int_0^T a(t)\frac{U_t\sigma(X_t) - Y_t\sigma(X_t)}{\sigma^2(X_t)}dW_t - \int_0^T \Big(\frac{a(t)Y_t}{\sigma'(X_t)}\Big)^2 dt - H_t \\ &= \frac{W_T^2}{x^2\sigma^2T^2} - \frac{W_T}{x^2\sigma T} - \frac{1}{x^2\sigma^2 T} - 0 \\ &= \frac{1}{x^2\sigma T}\Big(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\Big) \end{aligned} \tag{2.22}$$

These results are in accordance with the ones found in [13].

## 2.3 Optimal Weights

The weights found in the previous sections are not unique, which is readily seen from the weight found through the likelihood ratio method in equation (1.6) on page 7. Article [14] provided a way to determine how to find the optimal weights, where optimal means the weight with the least variance. As will be seen in the main proposition, any weights that are measurable with respect to $\mathcal{F}_T$ will be optimal, which includes all the weights found in the previous section.

Under the right conditions, such as in (L2.11), one can find:

$$\frac{\partial}{\partial x}\mathbb{E}[\Phi(X_T)] = \mathbb{E}\Big[\frac{\partial}{\partial x}\Phi(X_T)\Big] = \mathbb{E}[\Phi'(X_T)Y_T].$$

This is the precise solution. From now we adopt the notation used in [14] and consider the more general case:

$$\frac{\partial}{\partial \theta}\mathbb{E}[\Phi(F)] = \mathbb{E}\Big[\frac{\partial}{\partial \theta}\Phi(F)\Big] = \mathbb{E}[\Phi'(F)G].$$

The weights derived in previous sections are alternate ways to write this, i.e for some weight $\pi$ we have:

$$\mathbb{E}[\Phi'(F)G] = \mathbb{E}[\Phi(F)\pi]. \tag{2.23}$$

As mentioned above, the weight gained from the likelihood ratio method is one weight that satisfies (2.23), and the explicit expression was, as in equation (1.6):

$$\pi = \frac{\partial}{\partial \theta}\log f_\theta(F).$$

As we recall this is a mostly theoretical weight, since the density $f_\theta$ is rarely known explicitly. From [14], we know that for the the weight

$$\pi_0 = \frac{\partial}{\partial \theta}\log f_\theta\Big|_{\theta=0}(F),$$

equation (2.23) still holds, and with $\pi_0$ it is possible to define the class of all weights such that (2.23) is satisfied:

$$\mathscr{W} := \Big\{\pi \;\Big|\; \mathbb{E}[\pi \mid \mathcal{F}_T] = \pi_0\Big\}. \tag{2.24}$$

So, for any random variable $\pi$ such that $\mathbb{E}[\Phi(F)\pi] = \mathbb{E}[\Phi'(F)G]$, we have $\pi \in \mathscr{W}$.

In [14], the space of weights that is of interest is $\{\mathscr{W} \cap \mathbb{H}^1\}$, but as shown in (LA.8) on page 53, $\mathbb{H}^1 = \mathbb{D}_{1,2}$, so the space we will work with is:

$$\{\mathscr{W} \cap \mathbb{D}_{1,2}\}.$$

In the following we assume $F \in \mathbb{D}_{1,2}$ and that there exists some process $u_t$ such that the following equality holds:

$$\mathbb{E}\Big[\int_0^T D_tF \cdot u_t \;\Big|\; \mathcal{F}_T\Big] = \mathbb{E}[G \mid \mathcal{F}_T]. \tag{2.25}$$

For any $u_t$ satisfying (2.25), we can find a weight $\pi$:

$$\begin{aligned}
\mathbb{E}[\Phi'(F)G] &\overset{(*)}{=} \mathbb{E}\Big[\Phi'(F)\mathbb{E}[G \,|\, \mathcal{F}_T]\Big] \\
&\overset{(2.25)}{=} \mathbb{E}\Big[\Phi'(F)\mathbb{E}\Big[\int_0^T D_tFu_tdt \;\Big|\; \mathcal{F}_T\Big]\Big] \\
&\overset{(*)}{=} \mathbb{E}\Big[\Phi'(F)\int_0^T D_tFu_tdt\Big] \\
&= \mathbb{E}\Big[\int_0^T \Phi'(F)D_tFu_tdt\Big] \\
&\overset{(T2.7)}{=} \mathbb{E}\Big[\int_0^T D_t\Phi(F)u_tdt\Big] \\
&\overset{(T2.8)}{=} \mathbb{E}\Big[\Phi(F)\int_0^T u_t\delta W_t\Big] \\
&= \mathbb{E}\Big[\Phi(F)\delta(u_t)\Big] \\
&= \mathbb{E}[\Phi(F)\pi]
\end{aligned} \tag{2.26}$$

In the $(*)$ transitions, we applied the tower property for conditional expectations.

Using the set of processes $u_t$ that satisfy (2.25) we can define a new class of weights, which includes all the weights derived in the previous sections.

$$\big\{\pi = \delta(u_t) \in \mathbb{D}_{1,2} \mid u_t \text{ satisfies } (2.25)\big\}.$$

This set is equal to the set given in (2.24), as the following result shows.

**Proposition 2.23**
*Assuming $F \in \mathbb{D}_{1,2}$, the sets of weights introduced earlier are equal.*

$$\mathscr{W} \cap \mathbb{D}_{1,2} = \Big\{\pi = \delta(u_t) \in \mathbb{D}_{1,2} \mid u_t \textit{ satisfies } (2.25)\Big\}.$$

*Proof.*
Equality is proved by showing inclusion both ways.

$\supseteq$)
Assume $\pi = \delta(u_t) \in \mathbb{D}_{1,2}$ such that $u_t$ satisfies (2.25).

By assumption, $\pi \in \mathbb{D}_{1,2}$, so it remains to show that $\pi \in \mathscr{W}$, but this is also easily seen since $\mathscr{W}$ contains all weights that satisfies the equation (2.23) which $\pi = \delta(u_t)$ does through (2.26).

$\subseteq$)
For the converse we assume $\pi \in \mathscr{W} \cap \mathbb{D}_{1,2}$. Clearly $\pi \in \mathbb{D}_{1,2}$, so we must show $\pi = \delta(u_t)$ for some process $u_t$ that satisfies (2.25).

By choosing $\Phi = 1$, we get from (2.23):

$$\left.\begin{array}{c} \mathbb{E}[\Phi'(F)G] = \mathbb{E}[0] = 0 \\ \mathbb{E}[\Phi(F)\pi] = \mathbb{E}[\pi] \end{array}\right\} \implies \mathbb{E}[\pi] = 0. \tag{2.27}$$

Since $\pi \in \mathbb{D}_{1,2}$ we can apply the Clark-Ocone formula (T2.10).

$$\begin{aligned} \pi &\overset{(T2.10)}{=} \mathbb{E}[\pi] + \int_0^T \mathbb{E}[D_t\pi \mid \mathcal{F}_t] dW_t \\ &\overset{(2.27)}{=} \int_0^T \mathbb{E}[D_t\pi \mid \mathcal{F}_t] dW_t \\ &\overset{(T2.4)}{=} \int_0^T \mathbb{E}[D_t\pi \mid \mathcal{F}_t] \delta W_t \\ &= \delta(u_t), \end{aligned}$$

where we defined $u_t := \mathbb{E}[D_t\pi \,|\, \mathcal{F}_t]$ which is an adapted process with respect to $\mathcal{F}_t$, allowing us to apply Theorem 2.4.

By the tower property, now for any $\Phi$,

$$\mathbb{E}[\Phi'(F)G] = \mathbb{E}\Big[\Phi'(F)\mathbb{E}[G \,|\, \mathcal{F}_T]\Big]. \tag{2.28}$$

Since $\pi = \delta(u_t) \in \mathscr{W}$ by assumption, we know by the properties of $\mathscr{W}$ that (2.23) holds:

$$\mathbb{E}[\Phi'(F)G] = \mathbb{E}[\Phi(F)\pi],$$

and in addition, from the equalities in (2.26), working our way backwards:

$$\mathbb{E}[\Phi(F)\pi] = \mathbb{E}\left[\Phi'(F)\mathbb{E}\left[\int_0^T D_t F u_t dt \,\middle|\, \mathcal{F}_T\right]\right],$$

and by combining these equalities with (2.28), we get

$$\mathbb{E}\Big[\Phi'(F)\mathbb{E}[G \,|\, \mathcal{F}_T]\Big] = \mathbb{E}\left[\Phi'(F)\mathbb{E}\left[\int_0^T D_t F u_t dt \,\middle|\, \mathcal{F}_T\right]\right],$$

which means equation (2.25) is satisfied and proving the inclusion. ∎

The following is the main result in this section.

**Proposition 2.24**
*The weight $\pi_0$ yields the minimum variance over all $\pi \in \mathscr{W} \cap \mathbb{D}_{1,2}$ for the convex functional:*

$$\mathscr{V}(\pi) := \mathbb{E}\Big[\big|\Phi(F)\pi - \mathbb{E}[\Phi'(F)G]\big|^2\Big].$$

*Proof.*

$$\begin{aligned}
\mathscr{V}(\pi) &= \mathbb{E}\Big[\big|\Phi(F)\pi - \mathbb{E}[\Phi'(F)G]\big|^2\Big] \\
&= \mathbb{E}\Big[\Big(\Phi(F)[\pi \pm \pi_0] - \mathbb{E}[\Phi'(F)G]\Big)^2\Big] \\
&= \mathbb{E}\Big[\Big(\Phi(F)[\pi - \pi_0] + \Phi(F)\pi_0 - \mathbb{E}[\Phi'(F)G]\Big)^2\Big] \\
&\overset{(**)}{=} \mathbb{E}\big[\big(\Phi(F)[\pi - \pi_0]\big)^2\big] + \mathbb{E}\big[\big(\Phi(F)\pi_0 - \mathbb{E}[\Phi'(F)G]\big)^2\big] \\
&= \mathbb{E}\big[\Phi(F)^2[\pi - \pi_0]^2\big] + \mathscr{V}(\pi_0) \\
&\geq \mathscr{V}(\pi_0)
\end{aligned}$$

For any $\pi \in \mathscr{W} \cap \mathbb{D}_{1,2}$ we see that $\mathscr{V}(\pi) \geq \mathscr{V}(\pi_0)$, with equality if $\pi = \pi_0$, hence $\pi_0$ provides the minimal variance and is therefore the optimal choice.

To complete the proof we must verify step $(**)$. After multiplying the parenthesis, the additional term becomes 0, as we will show.

$$2\mathbb{E}\big[\big(\Phi(F)[\pi - \pi_0]\big)\big(\Phi(F)\pi_0 - \mathbb{E}[\Phi'(F)G]\big)\big].$$

Ignoring the factor 2, and using the tower property of conditional expectations:

$$\mathbb{E}\Big[\mathbb{E}\big[\big(\Phi(F)[\pi - \pi_0]\big)\big(\Phi(F)\pi_0 - \mathbb{E}[\Phi'(F)G]\big) \bigm| \mathcal{F}_T\big]\Big].$$

Now we can use that $\pi_0$, $\Phi(F)$ and $\mathbb{E}[\Phi'(F)G]$ are all $\mathcal{F}_T$-measurable.

$$\mathbb{E}\Big[\big(\Phi(F)\pi_0 - \mathbb{E}[\Phi'(F)G]\big) \cdot \underbrace{\mathbb{E}\big[\big(\Phi(F)[\pi - \pi_0]\big) \bigm| \mathcal{F}_T\big]}_{=0}\Big] = 0,$$

since:

$$\begin{aligned}
\mathbb{E}\big[\big(\Phi(F)[\pi - \pi_0]\big) \bigm| \mathcal{F}_T\big] &= \mathbb{E}\big[\Phi(F)\pi - \Phi(F)\pi_0 \bigm| \mathcal{F}_T\big] \\
&= \Phi(F)\mathbb{E}[\pi|\mathcal{F}_T] - \Phi(F)\pi_0 \\
&\overset{(2.24)}{=} \Phi(F)\pi_0 - \Phi(F)\pi_0 \\
&= 0
\end{aligned}$$

■

# Chapter 3

# Numerical Implementation of the BNS Model

This chapter will provide an example of how the Malliavin weights are calculated numerically for a stock price modelled by Lévy model. The model discussed is the Barndorff-Nielsen and Shephard model (BNS) which was introduced in [2].

The first section gives some background information on stochastic volatility models and proceeds to introduce the BNS model, with a brief discussion on its properties. The second section uses results from [4] and derives the Malliavin weights for $\Delta$ and $\Gamma$. The third section examines the technical details involved when the Malliavin weights are calculated numerically.

## 3.1 BNS: A Stochastic Volatility Model

As the theory of financial mathematics progresses, one important research topic is the construction of ever more realistic stochastic models that fit better to empirical data. One problem that has been of interest is how to model the volatility in the market, or more precisely the variation in the price of a stock over time, which is an important parameter that greatly affects model based quantities like option prices and Greeks.

In the Black-Scholes model the volatility is simply supposed to be constant, implying that the stock price varies roughly the same at all times, but this is not consistent with what is observed in reality.

The first attempt to rectify this was by introducing volatility as a function of time and stock price in the *local volatility models*, where the stochastic differential equations (from [9]) follow the dynamics:

$$dS_t = \mu S_t dt + \sigma(t, S_t) S_t dW_t.$$

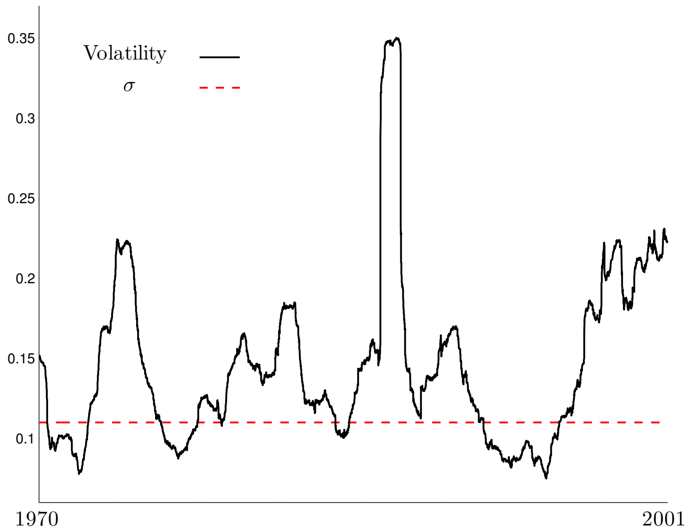


Figure 3.1: Historical volatility on the S&P500 from 1970-2001.

When combined with a time/price-dependant drift-function we have the general stochastic differential equations from Chapter 2, as given in (2.2).

Since the observed historical volatility varies in such an erratic way, the choice soon fell on modelling the volatility as a separate stochastic model, for instance as a bi-variate diffusion $(S_t, \sigma_t)$ driven by a two-dimensional Brownian motion, as seen in [9]. A general model in this case can be given by:

$$dS_t = \mu S_t dt + \sigma_t S_t dW_t^1,$$

with the stochastic volatility modelled by:

$$\sigma_t = f(Y_t), \quad dY_t = \alpha_t dt + \gamma_t dW_t^2.$$

The problem with this continuous stochastic model is that there are observed jumps in the volatility in the same way as for stock prices, and they often occur at the same time, e.g when the stock market crashes.

In Figure 3.1 there is a plot of the historical annual volatility from the S&P 500 market, which is recreated from a figure in [19]. As we see, the volatility has a tendency to increase rapidly or jump (the huge peak is the volatility following the 1987 stock market crash), and then fall back to some relative stable level, marked on the plot as the $\sigma$, as this could have been a choice for the constant volatility in the geometric Brownian motion model. The behaviour we see supports the idea that the level of randomness for the volatility jumps, like for a Lévy process, and steadily falls back down in a way that can be described by mean reversion.

A stochastic volatility model that is constructed in this way is the Barndorff-Nielsen and Shephard model, (BNS) which was introduced in [2].

## The Barndorff-Nielsen and Shephard Model (BNS)

The BNS model is a stochastic volatility, jump diffusion model where the squared volatility is modelled as a Lévy-driven positive Ornstein-Uhlenbeck process. If $\beta = -\frac{1}{2}$, $\sigma_t$ was constant and we removed $Z_t$ entirely, the model would be a geometric Brownian motion, so the BNS model can in a sense be viewed as a stochastic volatility extension. Following the general price model for Lévy processes as given in (1.10) on page 10, the BNS model is on the form

$$S_t = S_0 \exp\{X_t\},$$

where the Lévy process $X_t$ satisfies:

$$dX_t = (\mu + \beta\sigma_t^2)dt + \sigma_t dW_t + \rho dZ_{\lambda t}, \qquad X_0 = 0, \tag{3.1}$$

where $\rho \leq 0$, and the volatility follows the dynamics:

$$d\sigma_t^2 = -\lambda\sigma_t^2 dt + dZ_{\lambda t}, \qquad \sigma_0^2 > 0, \tag{3.2}$$

where $\lambda > 0$. The Lévy process, $Z_t$, assumed not to have a drift term, is usually called the background driving Lévy process (BDLP). $Z_t$ is a *subordinator*; a non-decreasing Lévy process, or more informally: the process $Z_t$ only has positive jumps. The path of the volatility $\sigma_t^2$ then moves by jumping up due to $Z_t$, and subsequently decays exponentially until it jumps again, imitating trends where new information causes a sudden upsurge in uncertainty, followed by a steady decrease to some relatively normal level.

In the BNS model the jumps by $Z_t$ affect both the stock price and the volatility at the same time through the negative parameter $\rho$. When the volatility rises, there is more risk associated with the stock and the price falls with a negative jump. The $\rho$ is said to be the *leverage effect*.

Note the Lévy process $Z_t$ has been chosen to follow a shifted time $\lambda t$, which is to ensure that the marginal distribution of $\sigma_t^2$ remains unchanged, no matter what $\lambda$ is chosen to be, as stated in [2].

The BNS model has been shown to be an arbitrage-free model, and since the market is not complete, there are several risk neutral measures. Under some of the risk neutral measures, the model ceases to be a proper BNS model since the Brownian motion and the Lévy process become dependent. To avoid this, the risk neutral measure that is used is assumed to be a *structure preserving* measure, in the sense that the model remains of the BNS type. Under the new measure the volatility remains unchanged as (3.2), whereas $dX_t$ now follows the dynamics:

$$dX_t = \big(r - \lambda\kappa(\rho) - \frac{1}{2}\sigma^2(t)\big)dt + \sigma(t)dW_t + \rho dZ_{\lambda t}, \tag{3.3}$$

where $\kappa(\rho)$ is the cumulant generating function under the risk neutral measure and $r > 0$ is the risk-less interest rate.

## 3.2 Malliavin Weights for the BNS model

Following [11] and [4] the Malliavin weights for $\Delta$ and $\Gamma$ will be derived for the BNS model. The BNS model does not follow the general stochastic differential equation in (2.2) used in section 2.2, so the corresponding Malliavin weight propositions cannot be used. The main difference is the presence of the subordinator $Z_t$, which is dealt with by taking the Malliavin derivative in the direction of the Brownian motion.

For the Wiener probability space $(\Omega^W, \mathcal{F}^W, Q^W)$ where $W_0 = 0$, $Q^W$-a.s, which is as used in Chapter 2 and the Lévy probability space $(\Omega^Z, \mathcal{F}^Z, Q^Z)$, the BNS model is modelled on the product of these spaces:

$$(\Omega, \mathcal{F}, Q) := (\Omega^W \otimes \Omega^Z, \mathcal{F}^W \otimes \mathcal{F}^Z, Q^W \otimes Q^Z).$$

The technical details are provided in [4], but an important consequence of this construction is that for any $\mathcal{F}^Z$-measurable random variable $F$, it follows that $D_t F = 0$ for the Malliavin derivative in the direction of the Brownian motion.

Denoting the initial value by $x$, and noting there is no dependence on $x$ in $X_t$, the first variation process becomes:

$$Y_t := \frac{\partial}{\partial x} S_T^x = \frac{\partial}{\partial x} x e^{X_T} = e^{X_T} = \frac{1}{x} S_T^x. \tag{3.4}$$

In Lemma A.10 on page 55, the additional equality is verified:

$$\frac{1}{x} S_T^x = \int_0^T \frac{a(t)}{x\sigma(t)} D_t S_T^x dt \tag{3.5}$$

### Delta

$$\begin{aligned}
\frac{\partial}{\partial x}\mathbb{E}[\Phi(S_T^x)] &\overset{(L2.11)}{=} \mathbb{E}\Big[\Phi'(S_T^x)\frac{\partial}{\partial x}S_T^x\Big] \\
&\overset{(3.4)}{=} \mathbb{E}\Big[\Phi'(S_T^x)\frac{1}{x}S_T^x\Big] \\
&\overset{(3.5)}{=} \mathbb{E}\Big[\int_0^T \Phi'(S_T^x) D_t S_T^x \frac{a(t)}{x\sigma(t)} dt\Big] \\
&\overset{(T2.7)}{=} \mathbb{E}\Big[\int_0^T D_t \Phi(S_T^x) \frac{a(t)}{x\sigma(t)} dt\Big] \\
&\overset{(T2.8)}{=} \mathbb{E}\Big[\Phi(S_T^x) \int_0^T \frac{a(t)}{x\sigma(t)} \delta W_t\Big] \\
&\overset{(T2.4)}{=} \mathbb{E}\Big[\Phi(S_T^x) \int_0^T \frac{a(t)}{x\sigma(t)} dW_t\Big]
\end{aligned}$$

## Gamma

We denote $\pi^{\Delta}$ as $F^x$, following the notation in [4]. The only dependence on $x$ in $F^x$ is the constant factor $x$ in the denominator, so $\frac{\partial}{\partial x}F^x = -\frac{1}{x}F^x$.

$$\begin{aligned}
\frac{\partial^2}{\partial x^2}\mathbb{E}[\Phi(S_T^x)] &= \frac{\partial}{\partial x}\mathbb{E}[\Phi(S_T^x)F^x] \\
&= \mathbb{E}\Big[F^x\frac{\partial}{\partial x}\Phi(S_T^x)\Big] + \mathbb{E}\Big[\Phi(S_T^x)\frac{\partial}{\partial x}F^x\Big] \\
&= \mathbb{E}\Big[F^x\frac{\partial}{\partial x}\Phi(S_T^x)\Big] - \frac{1}{x}\mathbb{E}[\Phi(S_T^x)F^x]
\end{aligned} \tag{3.6}$$

Calculating the first term in a similar way as when deriving the Malliavin weight for $\Delta$, but now with the additional factor $F^x$.

$$\begin{aligned}
\mathbb{E}\Big[F^x\frac{\partial}{\partial x}\Phi(S_T^x)\Big] &= \mathbb{E}\Big[\Phi'(S_T^x)\Big(\frac{\partial}{\partial x}S_T^x\Big)F^x\Big] \\
&\overset{(3.4)}{=} \mathbb{E}\Big[\Phi'(S_T^x)\frac{1}{x}S_T^xF^x\Big] \\
&\overset{(3.5)}{=} \mathbb{E}\Big[\int_0^T \Phi'(S_T^x)D_tS_T^x\frac{a(t)}{x\sigma(t)}F^x dt\Big] \\
&\overset{(T2.7)}{=} \mathbb{E}\Big[\int_0^T D_t\Phi(S_T^x)\frac{a(t)}{x\sigma(t)}F^x dt\Big] \\
&\overset{(T2.8)}{=} \mathbb{E}\Big[\Phi(S_T^x)\int_0^T \frac{a(t)}{x\sigma(t)}F^x\delta W_t\Big] \\
&= \mathbb{E}\Big[\Phi(S_T^x)\delta\Big(\frac{a(t)}{x\sigma(t)}\cdot F^x\Big)\Big]
\end{aligned} \tag{3.7}$$

Defining $u(t) := \frac{a(t)}{x\sigma(t)}$. Since there is no dependence on the Wiener process in $u(t)$ we can regard it as a deterministic function with respect to the Malliavin derivative in the Wiener direction:

$$D_tF^x = D_t\Big(\int_0^T u(s)dW_s\Big) = u(t) = \frac{a(t)}{x\sigma(t)}.$$

Applying integration by parts to the Skorohod integral:

$$\begin{aligned}
\delta\Big(\frac{a(t)}{x\sigma(t)}\cdot F^x\Big) &\overset{(T2.9)}{=} F^x\delta(u(t)) - \int_0^T D_tF^xu(t)dt \\
&\overset{(T2.4)}{=} F^x\int_0^T u(s)dW_s - \int_0^T D_tF^xu(t)dt \\
&= (F^x)^2 - \int_0^T u^2(t)dt
\end{aligned} \tag{3.8}$$

Collecting the calculations:

$$\begin{aligned}
\frac{\partial^2}{\partial x^2}\mathbb{E}[\Phi(S_T^x)] &\overset{(3.6)}{=} \mathbb{E}\Big[F^x \frac{\partial}{\partial x}\Phi(S_T^x)\Big] - \frac{1}{x}\mathbb{E}[\Phi(S_T^x)F^x] \\
&\overset{(3.7)}{=} \mathbb{E}\Big[\Phi(S_T^x)\delta\Big(\frac{a(t)}{x\sigma(t)}\cdot F^x\Big)\Big] - \frac{1}{x}\mathbb{E}[\Phi(S_T^x)F^x] \\
&\overset{(3.8)}{=} \mathbb{E}\Big[\Phi(S_T^x)\Big((F^x)^2 - \int_0^T u^2(t)dt - \frac{1}{x}F^x\Big)\Big] \\
&= \mathbb{E}\Big[\Phi(S_T^x)\Big((\pi^\Delta)^2 - \frac{1}{x}\pi^\Delta - \int_0^T u^2(t)dt\Big)\Big]
\end{aligned}$$

We have found the Malliavin weight: $\pi^\Gamma = ((\pi^\Delta)^2 - \frac{1}{x}\pi^\Delta - \int_0^T u^2(t)dt)$.

The Lévy process $Z_{\lambda t}$ was differentiated away through the use of equation (3.5), and from there the calculations follow the same, general approach as the calculations for the continuous case from Chapter 2. Deriving the Malliavin weights through the Malliavin derivative in the Wiener direction can be used to find the Malliavin weights in all jump diffusion models, as in e.g [10] among others.

Writing out the full expressions for these weights with the choice $a(t) = \frac{1}{T}$, we have:

$$\begin{aligned}
\pi^\Delta_{BNS} &= \int_0^T \frac{1}{x\sigma(t)T}dW_t \\
\pi^\Gamma_{BNS} &= (\pi^\Delta_{BNS})^2 - \frac{1}{x}\pi^\Delta_{BNS} - \int_0^T \frac{1}{x^2\sigma^2(t)T^2}dt
\end{aligned}$$

Recalling from the Example in Chapter 2, the Malliavin weights for the Black-Scholes model were found to be in equations (2.18) and (2.22):

$$\begin{aligned}
\pi^\Delta_{BS} &= \frac{W_T}{x\sigma T} = \int_0^T \frac{1}{x\sigma T}dW_t \\
\pi^\Gamma_{BS} &= (\pi^\Delta_{BS})^2 - \frac{1}{x}\pi^\Delta_{BS} - \int_0^T \frac{1}{x^2\sigma^2 T^2}dt
\end{aligned}$$

We clearly see the connection between the Black-Scholes model and the stochastic volatility extension of the BNS model, as the weights are precisely the same, just with the constant $\sigma$ exchanged with the random variable $\sigma(t)$.

## 3.3 Numerical Simulation

In simulating the BNS-model, the main challenge is simulating the stochastic volatility. From [4] we have the explicit solution to the stochastic differential equation from (3.2) given by:

$$\sigma^2(t) = \sigma^2(0)e^{-\lambda t} + \int_0^t e^{\lambda(s-t)} dZ_{\lambda s}, \qquad \sigma^2(0) > 0.$$

From [19] or [4], we can rewrite this expression by transferring the time shifting factor $\lambda$ to the integration limit in the Lévy integral.

$$\sigma^2(t) = \sigma^2(0)e^{-\lambda t} + e^{-\lambda t} \int_0^{\lambda t} e^s dZ_s, \qquad \sigma^2(0) > 0. \tag{3.9}$$

To simulate this expression we need to simulate the Lévy integral. This can be done through the series representation of the Lévy integral as given in [19]:

$$\int_0^t f(s) dZ_s = \sum_{i=1}^{\infty} W^{-1}\Big(\frac{a_i}{t}\Big) f(t \cdot u_i) \mathbf{1}_{\{u_i < t\}}, \tag{3.10}$$

where the equality is in law, where the sequence of random variables $\{a_i\}$ are arrival times for a Poisson process with intensity 1, and $\{u_i\}$ are uniform $U[0,1]$ variables. Furthermore, the function $W^{-1}(\cdot)$ denotes the inverse of the tail mass function, which is defined as:

$$W^+(x) = \int_x^\infty \nu(dy),$$

where $\nu(dy)$ is the Lévy measure for the Lévy process $Z_t$.

The jumps will henceforth be modelled by the inverse Gaussian distribution IG$(\delta, \gamma)$, meaning the process $Z_t$ is an inverse Gaussian process. From [19] the Lévy measure for the IG$(\delta, \gamma)$-distribution is given by:

$$\nu(dx) = \frac{\delta}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp\Big\{ -\frac{1}{2}\gamma^2 x \Big\} dx,$$

from which we can extract the Lévy density $u(x)$:

$$u(x) = \frac{\delta}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp\Big\{ -\frac{1}{2}\gamma^2 x \Big\}.$$

The tail mass function is related to the Lévy density by: $W^+(x) = x \cdot u(x)$, which yields:

$$W^+(x) = xu(x) = x\Big(\frac{\delta}{\sqrt{2\pi}} x^{-\frac{3}{2}} \exp\Big\{ -\frac{1}{2}\gamma^2 x \Big\}\Big) = \frac{\delta}{\sqrt{2\pi}} x^{-\frac{1}{2}} \exp\Big\{ -\frac{1}{2}\gamma^2 x \Big\},$$

which is in accordance with [2]. Matlab code implemented for this function is included in `Wplus.m` on page 58. This expression for the tail mass function is not analytically invertible, so instead we must find the inverse numerically. The method will be based on the following definition of the inverse tail mass function:

$$W^{-1}(x) := \inf\big\{y > 0 \mid W^{+}(y) \leq x\big\}.$$

The numerical inverse $W^{-1}$ has been implemented in Matlab and is included as `Winv.m` on page 58.

For the simulations both the Lévy integral and the Lévy process are needed for $\sigma^2(t)$ and $X_t$, respectively, so for the BNS-model, equation (3.10) becomes:

$$\int_0^t f(s)dZ_s = \int_0^t \exp(s)dZ_s \approx \sum_{i=1}^{N_\varepsilon} W^{-1}\Big(\frac{a_i}{t}\Big)\exp(t\cdot u_i)\mathbf{1}_{\{u_i<t\}},$$

and

$$\int_0^t f(s)dZ_s = \int_0^t 1dZ_s = Z_t \approx \sum_{i=1}^{N_\varepsilon} W^{-1}\Big(\frac{a_i}{t}\Big)\mathbf{1}_{\{u_i<t\}}.$$

which are both simulated in `subord.m` on page 59. The sums cannot be calculated with an infinite amount of terms for obvious reasons, so the sums are taken to the finite limit $N_\varepsilon$, which is suggested in [15] and is defined as:

$$N_\varepsilon := \inf\big\{j \in \mathbb{N} \mid W^{-1}\Big(\frac{a_i}{t}\Big) \leq \frac{1}{N}\big\},$$

where $1/N$ is the first step in the discretized interval used for `Winv`; in other words we are calculating up to the level of precision enforced by the discretization.

Since the Lévy process used in $X_t$ is an IG-process, we can simulate the process directly through `IGproc.m` on page 61 and compare with the series representation. This is done in figure 3.2 on the facing page, with some randomly chosen parameters $\delta = 5$ and $\gamma = 20$. The trajectories are similar and give a good indication that the code has been correctly implemented; the path of the process generated from the series representation may vary slightly from the other paths as it is a much cruder approximation to the IG-process. The main advantage of the series representation is that it is easy to use when simulating Lévy integrals, which makes up for some of the approximation error.

When the Lévy integral has been simulated, it is simple to implement the volatility process $\sigma^2(t)$ and the corresponding BNS path. In figure 3.3 on page 43 there are plots of $Z_t$ and the volatility process $\sigma^2(t)$ with initial value $\sigma^2(0) = 0.2$. The jumps for $\sigma^2(t)$ which are modelled through the Lévy integral in equation (3.10) are seen to follow the same general pattern as the underlying Lévy process $Z_t$, and we clearly see the effect of the mean reversion.

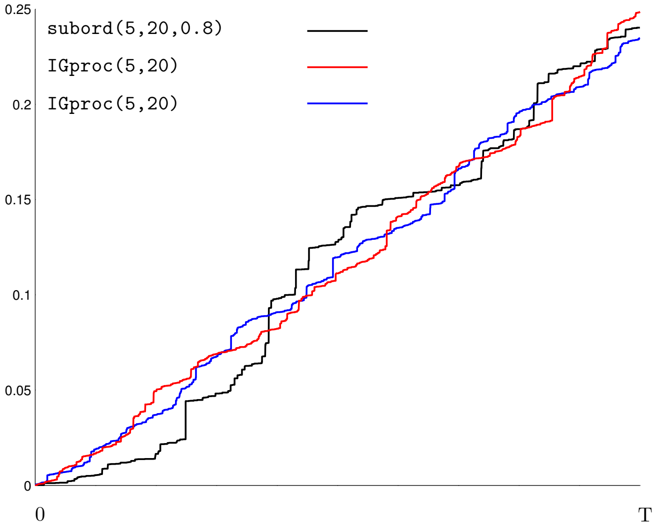


Figure 3.2: Comparison of IG-processes with $\delta = 5$ and $\gamma = 20$.

Recalling the dynamics for $X_t$ under the risk neutral probability measure as in equation (3.3):

$$dX_t = \big(r - \lambda\kappa(\rho) - \frac{1}{2}\sigma^2(t)\big)dt + \sigma(t)dW_t + \rho dZ_{\lambda t}.$$

All terms have been accounted for except the cumulant generating function, which for an inverse Gaussian distribution has the form:

$$\kappa(z) = \frac{z\delta}{\sqrt{\gamma - 2z}},$$

as seen in [4]. Simulation of the BNS process is done in `bns.m` on page 60.

Two paths of the BNS model are shown in figure 3.4, demonstrating the influence of the leverage effect for different values of $\rho$. In this simulation the parameters were chosen to be $\delta = 5$ and $\gamma = 20$ (which are the same parameters used earlier), so the jumps - as we can see in the plot of $\sigma^2(t)$ - are frequent and small, but their overall effect on the BNS model is apparent from the plot of the BNS paths. To avoid some technicalities, the time shift $\lambda$ is chosen to be 1.

The final part in this numerical implementation is programming a way to calculate the Malliavin weights for the BNS model, but as can be seen from the weights derived in the previous section this only includes calculating the Ito integral:

$$\int_0^T \frac{1}{\sigma(t)} dW_t,$$

and the deterministic integral:

$$\int_0^T \frac{1}{\sigma^2(t)}dt.$$

Calculating the Ito integral is done by a finite sum approximation based on the definition of the Ito integral. For the partition $0 = t_1 \leq t_2 \leq \ldots \leq t_{n-1} \leq t_n = T$, we have:

$$\int_0^T \frac{1}{\sigma(t)}dW_t \approx \sum_{i=1}^{n} \frac{1}{\sigma(t_i)}\big(W_{t_{i+1}} - W_{t_i}\big) \implies$$

$$\pi^{\Delta} = \int_0^T \frac{1}{x\sigma(t)T}dW_t \approx \frac{1}{xT}\sum_{i=1}^{n-1} \frac{1}{\sigma(t_i)}\big(W_{t_{i+1}} - W_{t_i}\big)$$

The deterministic integral can be calculated using the standard trapezoidal rule. Assuming all the time steps are equal:

$$\int_0^T \frac{1}{\sigma^2(t)}dt \;\approx\; \sum_{i=1}^{n-1} \frac{1}{2}\Big(\frac{1}{\sigma^2(t_{i+1})} + \frac{1}{\sigma^2(t_i)}\Big) \cdot \frac{T}{n} \implies$$

$$\begin{aligned}
\pi^{\Gamma} &= (\pi^{\Delta})^2 - \frac{1}{x}\pi^{\Delta} - \int_0^T \frac{1}{x^2\sigma^2(t)T^2}dt \\
&\approx (\pi^{\Delta})^2 - \frac{1}{x}\pi^{\Delta} - \frac{1}{x^2T^2}\frac{T}{2n}\sum_{i=1}^{n-1}\Big(\frac{1}{\sigma^2(t_{i+1})} + \frac{1}{\sigma^2(t_i)}\Big) \\
&= (\pi^{\Delta})^2 - \frac{1}{x}\pi^{\Delta} - \frac{1}{2nx^2T}\sum_{i=1}^{n-1}\Big(\frac{1}{\sigma^2(t_{i+1})} + \frac{1}{\sigma^2(t_i)}\Big) \\
&= (\pi^{\Delta})^2 - \frac{1}{x}\pi^{\Delta} - \frac{1}{2nx^2T}\bigg(\frac{1}{\sigma^2(t_n)} + \frac{1}{\sigma^2(t_1)} + \sum_{i=2}^{n-1}\frac{2}{\sigma^2(t_i)}\bigg).
\end{aligned}$$

Since the function $\sigma^2(t)$ is discontinuous, there will be some errors in the numerical approximations. The error could be reduced by making finer grids, or eliminated by calculating the integrals over the piecewise continuous parts. Since this is just intended as a quick description of the numerical method, the technical details of these approaches are left out. A suggested Matlab program to calculate the Malliavin weights is included in `bnsWeights.m` on page 61, where the expiration time $T$ is assumed to be 1.

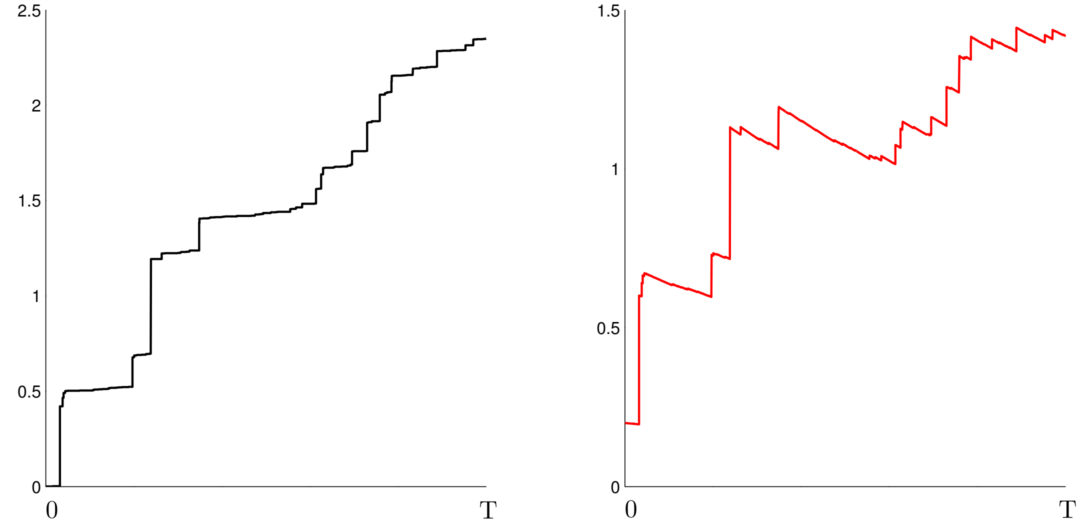


Figure 3.3: Processes $Z_t$ and $\sigma^2(t)$ for $\delta = 5$ and $\gamma = 0.2$.

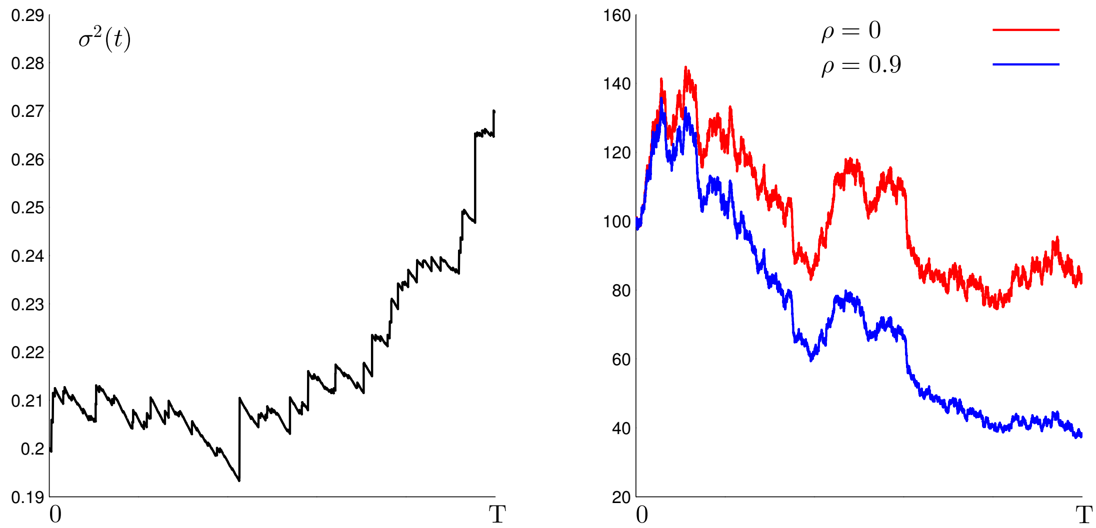


Figure 3.4: $\sigma^2(t)$ and corresponding BNS path with different leverages.

# Chapter 4

# Conclusion

## 4.1 Summary

The Greeks are measures of considerable importance for managing risk associated with financial derivatives, and improving and extending the theory is of great interest to applied financial mathematics and financial engineering.

Fournié et al. introduced in their 1999 article [13] a method for calculating the derivatives of the pay-off functions using Malliavin weights as seen in Chapter 2, which provided an improvement when compared to alternative methods such as the finite difference method and likelihood ratio method as discussed in Chapter 1. The improvement is first and foremost that of a wider range of applicability; for calculating the delta for a call option, the finite difference method would be preferable as it is easier to implement and gives the same convergence rate.

In Lemma 2.24 from section 2.3 it was shown that for all possible Malliavin differentiable weights, the additional condition of $\mathcal{F}_T$-measurability means the weight is optimal in the sense of minimal variance. These conditions apply to all the weights derived in Chapter 2.

As noted in section 1.3, continuous models are not able to give a realistic description of how real world stocks and commodities behave, a topic that is covered more extensively in [9] or [19], which leads to the introduction of Lévy processes. Finding the Malliavin weights in the discontinuous case is an extension of the methods from Chapter 2, and for the jump diffusion model it is a simple matter of taking the Malliavin derivative in the direction of the Brownian motion as in Chapter 3. From thereon there are few changes from the continuous case.

An important drawback with the calculation of the Greeks in jump diffusion models, is when the jumps are modelled by Lévy processes that are more complicated than the compound Poisson process; the numerical calculations become too involved which makes any real world applications difficult.

## 4.2 Possible Extensions

Due to time constraints a number of discussions that were intended to appear in the thesis were left out, as well as a few interesting related subjects discovered during the research that was conducted. A few of these topics are presented here as possible extensions to this thesis.

- The focus of this thesis has been learning and applying Malliavin calculus. Because of this, a proper mathematical treatment of Lévy processes has been left out, most notably the details around the argument used to show how to derive the Malliavin weights for jump diffusion models. Given the importance of Lévy models, a detailed understanding of the processes are essential, and in the framework of Greeks, so is the Malliavin calculus for Lévy processes.

- A variance reduction technique called *localised Malliavin formula* was introduced in [13], which for an e.g European call option involves localising the Malliavin weights around the strike value $K$. This was shown to significantly increase the convergence rate, improving the method further. In fact, without this vital extension, the finite different method usually outperformed the method using the Malliavin weights.

- Article [14] introduces a new application of Malliavin calculus to the representation of conditional expectations. By using the integration by parts formula (T2.9) they find an expression for the conditional expectation that is possible to simulate directly. This opens for the possibility of calculating the Greeks for more advanced financial derivatives such as American options. The authors of [13] and [14] stated they considered this to be one of the most promising future applications of Malliavin calculus to numerical finance.

- While the Malliavin weights for the Greeks carry over easily from the continuous case to jump diffusions, this is not the case with infinite activity Lévy models (which were discussed in Chapter 1). In fact, finding the Greeks for such models remains an unsolved problem.

  One approach that could provide a rough estimate is by considering an infinite activity Lévy model of the normal inverse Gaussian (NIG) type, based on the NIG distribution introduced in [1], and approximate it by a geometric Brownian motion. The drift and volatility for the geometric Brownian motion are chosen to be the expectation and variance of the NIG distribution, and for a simulation of the NIG Lévy model, a simulation of a geometric Brownian approximation can be "fitted" to follow the NIG process.

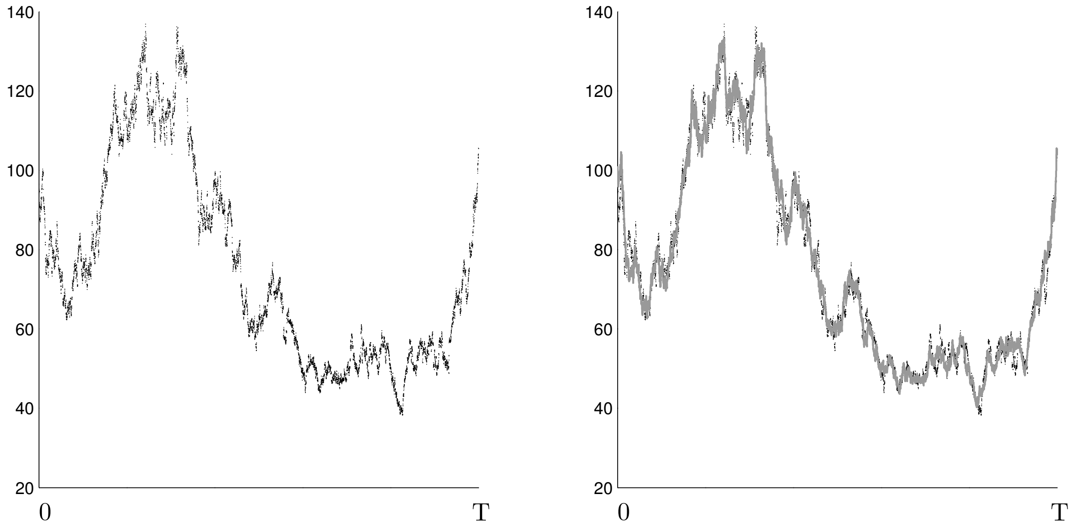


Figure 4.1: NIG Process and GBM Approximation.

An illustration of this is shown in figure 4.1, where the left side shows a "well behaved" NIG process in the sense that it does not jump around too much (like for instance the NIG process depicted in figure 1.2 on page 12) and a simulation of a geometric Brownian motion approximation. The geometric Brownian motion was simulated by first choosing 50 evenly spaced points along the NIG trajectory and repeatedly simulating a geometric Brownian motion path between these points until it hits within a predefined limit $\varepsilon > 0$ from the NIG process. The run time of such an algorithm is surprisingly short, and the resulting path is a properly simulated geometric Brownian motion. For less well behaved NIG processes it is often possible to get a close approximation by using jump diffusion models.

Now we have two closely related objects where we know how to compute the Greeks for one of them, the problem being that the Malliavin weight for the geometric Brownian motion, as derived for the e.g delta in equation (2.18), depends completely on the geometric Brownian motion-model. If there exists a Malliavin weight for the NIG process, it should likewise be at least partly based on the distribution parameters and starting value which are all known. Since these values are also used in the geometric Brownian motion, this opens for the possibility that there is some algebraic relationship between the weights, which again means the Malliavin weight for the geometric Brownian motion can serve as a very crude approximation.

There is a lot of speculation and a lot of unanswered questions in this approach. Regrettably there was no time to pursue it, so it is listed as a possible extension.

# Appendix A

## A.1 Additonal Results

**Lemma A.1**
*Assuming $V(s,x)$ is a two time differentiable function, we have:*

$$dV(t,S_t) = \frac{\partial V}{\partial t}(t,S_t)dt + \frac{\partial V}{\partial x}(t,S_t)dS_t + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2}(t,S_t)dt$$

*where $dS_t$ is given by (1.1).*

*Proof.*
By a direct application of Ito's formula:

$$dV(t,S_t) = \frac{\partial V}{\partial t}(t,S_t)dt + \frac{\partial V}{\partial x}(t,S_t)dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial x^2}(t,S_t)(dS_t)^2 \tag{A.1}$$

Finding the square of (1.1):

$$\begin{aligned}(dS_t)^2 &= \big(\mu S_t dt + \sigma S_t dW_t\big)^2 \\ &= \mu^2 S_t^2 (dt)^2 + \sigma^2 S_t^2 (dW_t)^2 + 2\mu\sigma S_t^2 dt dW_t\end{aligned}$$

Using the standard stochastic calculus rules: $(dt)^2 = dtdW_t = 0$ and $(dW_t)^2 = dt$

$$= \sigma^2 S_t^2 dt. \tag{A.2}$$

Substituting this back into (A.1):

$$dV(t,S_t) = \frac{\partial V}{\partial t}(t,S_t)dt + \frac{\partial V}{\partial x}(t,S_t)dS_t + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial x^2}(t,S_t)dt,$$

and the equality is proved. ∎

**Lemma A.2** (Solution for Geometric Brownian motion)
*The solution to the stochastic differential equation*

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

*for* $\mu, \sigma \in \mathbb{R}$, *is:*

$$S_t = S_0 \exp\Big\{\Big(\mu - \frac{1}{2}\sigma^2\Big)t + \sigma W_t\Big\}.$$

*Proof.*
Defining the function $f(t,x) := \log(x)$ and calculating the partial derivatives:

$$\frac{\partial f}{\partial t}(t,x) = 0, \qquad \frac{\partial f}{\partial x}(t,x) = \frac{1}{x}, \qquad \frac{\partial^2 f}{\partial x^2}(t,x) = -\frac{1}{x^2}.$$

By the Ito formula:

$$\begin{aligned} d\big[\log S_t\big] &= \frac{\partial f}{\partial t}(t,S_t)dt + \frac{\partial f}{\partial x}(t,S_t)dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,S_t)(dS_t)^2 \\ &\overset{(A.2)}{=} 0 + \frac{1}{S_t}\big(\mu S_t dt + \sigma S_t dW_t\big) - \frac{1}{2}\frac{1}{S_t^2}(\sigma^2 S_t^2 dt) \\ &= \mu dt + \sigma dW_t - \frac{1}{2}\sigma^2 dt \\ &= \Big(\mu - \frac{1}{2}\sigma^2\Big)dt + \sigma dW_t \end{aligned}$$

Writing equation on integral form and calculating.

$$\begin{aligned} \log S_t &= \log S_0 + \int_0^t \Big(\mu - \frac{1}{2}\sigma^2\Big)ds + \int_0^t \sigma dW_s \\ &= \log S_0 + \Big(\mu - \frac{1}{2}\sigma^2\Big)t + \sigma W_t \end{aligned}$$

Applying the exponential function to both sides.

$$\begin{aligned} S_t &= \exp\Big\{\log S_0 + \Big(\mu - \frac{1}{2}\sigma^2\Big)t + \sigma W_t\Big\} \\ &= S_0 \exp\Big\{\Big(\mu - \frac{1}{2}\sigma^2\Big)t + \sigma W_t\Big\}. \end{aligned}$$

■

**Lemma A.3**
*For the exponential martingale,*

$$Z_t = \exp\Big\{ \int_0^t u(s,\omega) dW_s - \frac{1}{2}\int_0^t u^2(s,\omega) ds \Big\},$$

*we can write $Z_t$ as an Ito process with:*

$$dZ_t = Z_t u(t,\omega) dW_t,$$

*and*

$$Z_T = 1 + \int_0^T Z_t u(t,\omega) dW_t.$$

*Proof.*
We define

$$M_t := \int_0^t u(s,\omega) dW_s - \frac{1}{2}\int_0^t u^2(s,\omega) ds$$

so

$$dM_t = u(t,\omega) dW_t - \frac{1}{2} u^2(t,\omega) dt, \qquad M_0 = 0,$$

and $Z_T = \exp\{M_T\}$. We find the partial derivatives of $g(t,x) = e^x$,

$$\frac{\partial g}{\partial t} = 0, \quad \frac{\partial g}{\partial x} = e^x, \quad \frac{\partial^2 g}{\partial x^2} = e^x,$$

and apply Ito's formula to $g(t, M_t) = e^{M_t} = Z_t$:

$$dZ_t = 0 + e^{M_t} dM_t + \frac{1}{2} e^{M_t} (dM_t)^2. \tag{A.3}$$

Since $dW_t^2 = dt$ and $dt^2 = dW_t dt = dt dW_t = 0$, we get:

$$(dM_t)^2 = u^2(t,\omega) dt.$$

Substituting $(dM_t)^2$, $dM_t$ and $e^{M_t} = Z_t$ into (A.3) we find:

$$\begin{aligned} dZ_t &= Z_t\Big(u(t,\omega) dW_t - \frac{1}{2} u^2(t,\omega) dt\Big) + \frac{1}{2} Z_t u^2(t,\omega) dt \\ &= Z_t u(t,\omega) dW_t \cancel{- \frac{1}{2} Z_t u^2(t,\omega) dt} + \cancel{\frac{1}{2} Z_t u^2(t,\omega) dt} \\ &= Z_t u(t,\omega) dW_t \end{aligned}$$

Writing on integral form, and using $Z_0 = e^{M_0} = e^0 = 1$:

$$\begin{aligned} Z_T &= \mathbb{E}[Z_0] + \int_0^T Z_t u(t,\omega) dW_t \\ &= 1 + \int_0^T Z_t u(t,\omega) dW_t \end{aligned}$$

■

The following two Definitions, Proposition A.6 and Lemma A.7 are supporting results for Lemma A.8 which is a result used in the main text.

**Definition A.4**
*As defined in [14],* $\mathbb{H}^1$ *is the set of* $L^2(P)$*-variables* $F$ *such that* $D_t F \in L^2(P \times \lambda)$*:*

$$\mathbb{H}^1 := \left\{ F \in L^2(P) \;\middle|\; \|D_t F\|^2_{L^2(P\times\lambda)} < \infty \right\}$$

*or in short,* $F \in \mathbb{H}^1$ *if* $\|D_t F\|^2_{L^2(P\times\lambda)} < \infty$.

**Definition A.5**
*By the definition of the Malliavin differentiable random variables (D2.5), we have that* $F \in \mathbb{D}_{1,2}$ *if*

$$\|F\|^2_{\mathbb{D}_{1,2}} := \sum_{n=1}^{\infty} n n! \|f_n\|^2_{L^2[0,T]^n} < \infty.$$

**Proposition A.6**
*Assuming* $g$ *and* $h$ *are symmetric, square integrable functions,*

$$\mathbb{E}[I_n(g) I_m(h)] = \begin{cases} 0 & n \neq m \\ n!(g,h)_{L^2([0,T]^n)} & n = m \end{cases}$$

*Proof.*
Proposition 1.4 in [11]. ■

**Lemma A.7**

$$\mathbb{E}[(D_t F)^2] = \sum_{n=1}^{\infty} n^2 (n-1)! \|f_n(\cdot,t)\|^2_{L^2([0,T]^{n-1})}$$

*Proof.*

$$\begin{aligned}
\mathbb{E}[(D_t F)^2] &\overset{(D2.6)}{=} \mathbb{E}\Big[\Big(\sum_{n=1}^{\infty} n I_{n-1}\big[f_n(\cdot,t)\big]\Big)^2\Big] \\
&= \mathbb{E}\Big[\sum_{n=1}^{\infty} n^2 I^2_{n-1}\big[f_n(\cdot,t)\big] + \sum_{n\neq k} nk I_{n-1}\big[f_n(\cdot,t)\big] I_{k-1}\big[f_k(\cdot,t)\big]\Big] \\
&= \sum_{n=1}^{\infty} n^2 \mathbb{E}\Big[I^2_{n-1}\big[f_n(\cdot,t)\big]\Big] + \sum_{n\neq k} nk \underbrace{\mathbb{E}\Big[I_{n-1}\big[f_n(\cdot,t)\big] I_{k-1}\big[f_k(\cdot,t)\big]\Big]}_{=0 \text{ by } (PA.6)} \\
&\overset{(PA.6)}{=} \sum_{n=1}^{\infty} n^2 (n-1)! \|f_n(\cdot,t)\|^2_{L^2([0,T]^{n-1})}
\end{aligned}$$

■

The following result, which relies on the previous definitions and supporting lemmas, gives a relation between a space $\mathbb{H}^1$ defined in [14] and the space $\mathbb{D}_{1,2}$, which was introduced in section 2.1.

**Lemma A.8**

$$\mathbb{H}^1 = \mathbb{D}_{1,2}$$

*Proof.*
We first note that we have the following equalities:

$$\begin{aligned}
\|D_t F\|^2_{L^2(P\times\lambda)} &= \mathbb{E}\Big[\int_0^T (D_t F)^2 dt\Big] \\
&= \int_0^T \mathbb{E}[(D_t F)^2] dt \\
&\overset{(LA.7)}{=} \int_0^T \sum_{n=1}^{\infty} n^2(n-1)! \|f_n(\cdot,t)\|^2_{L^2([0,T]^{n-1})} dt \\
&= \sum_{n=1}^{\infty} n n! \int_0^T \|f_n(\cdot,t)\|^2_{L^2([0,T]^{n-1})} dt \\
&= \sum_{n=1}^{\infty} n n! \|f_n\|^2_{L^2([0,T]^n)} \\
&= \|F\|^2_{\mathbb{D}_{1,2}}
\end{aligned}$$

Since these are equal we get the equivalence:

$$\|D_t F\|^2_{L^2(P\times\lambda)} < \infty \iff \|F\|^2_{\mathbb{D}_{1,2}} < \infty.$$

From (DA.4) and (DA.5), we see that the membership requirements for $\mathbb{H}^1$ and $\mathbb{D}_{1,2}$ are equivalent, hence the sets are equal. ■

Not really a Lemma, just a verification that the random variable studied throughout this paper satisfies Lemma 2.11 for $\theta = x$.

**Lemma A.9**
*$F^\theta = \Phi(X_T^x) \in \mathbb{R}$ satisfies the conditions in Lemma 2.11 when $\theta = x$, i.e the mapping $x \mapsto \Phi(X_T^x)$, for a fixed $T$, is continuously differentiable, and*

$$\mathbb{E}\Big[\sup_{x\in[a,b]}\Big|\frac{\partial \Phi(X_T^x)}{\partial x}\Big|\Big] < \infty. \tag{A.4}$$

*Proof.*
We assume $x \in [0, b]$ where $b < \infty$. The partial derivative of $\Phi(X_T^x)$ with respect to $x$, is:

$$\frac{\partial}{\partial x}\Phi(X_T^x) = \Phi'(X_T^x)\frac{\partial}{\partial x}X_T^x = \Phi'(X_T^x)Y_T, \tag{A.5}$$

where $Y_T$ is the first variation process as given in equation (2.5) on page 17. Since the functions $\mu(\cdot)$ and $\sigma(\cdot)$ are Lipschitz continuous:

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leq D|x - y|,$$

for a finite $D$, and $\Phi(\cdot)$ has a bounded derivative, we have $\Phi'(X_T^x) < \infty$ for all $\omega \in \Omega$. The functions $\mu$ and $\sigma$ are also assumed to have bounded derivatives, which means $Y_T < \infty$ for all $\omega \in \Omega$, hence their product is finite and by going backwards in (A.5), so is $\frac{\partial}{\partial x}\Phi(X_T^x)$.

The finiteness is true as long as $x$ is finite, or $x \in [0, b]$, so equation (A.4) is verified. ■

**Lemma A.10**
*For the BNS model:*

$$S_T^x = xe^{X_T} = x\exp\Big\{X_0 + \int_0^T \big(r - \lambda\kappa(\rho) - \frac{1}{2}\sigma(t)^2\big)dt + \int_0^T \sigma(s)dW_s + \int_0^T \rho dZ_{\lambda t}\Big\},$$

*the Malliavin derivative in the Wiener direction is:*

$$D_t S_T^x = S_T^x \sigma(t). \tag{A.6}$$

*By this relation, it follows that:*

$$\frac{1}{x}S_T^x = \int_0^T \frac{a(t)}{x\sigma(t)} D_t S_T^x dt. \tag{A.7}$$

*Proof.*
First part: verifying equation (A.6).
We denote $S_T^x$ as a function $f$ of $X_T$: $S_T^x = f(X_T)$ where $f(y) = xe^y$ with $\frac{\partial}{\partial y} f(y) = xe^y = f(y)$. By the chain rule:

$$D_t S_T^x = D_t f(X_T) \overset{(T2.7)}{=} f'(X_T) D_t X_T = S_T^x D_t X_T.$$

Since the Malliavin derivative is a linear operator, each term in $X_T$ is evaluated individually. The Malliavin derivative is in the Wiener direction, so the initial value $X_0$, the deterministic $dt$-integral and the Lévy integral $dZ_{\lambda t}$ all become 0. Hence the equation follows by considering the $dW_s$-integral as a Wiener-Ito chaos expansion, or from the *fundamental theorem of stochastic calculus* as formulated in [11] (where $\sigma(\cdot)$ is regarded as deterministic with respect to the Wiener process):

$$D_t S_T^x = S_T^x D_t X_T = S_T^x \cdot D_t\Big(\int_0^T \sigma(s)dW_s\Big) = S_T^x \sigma(t).$$

Second part: verifying equation (A.7).
By choosing $a(t) \in \mathcal{A}$, i.e any function $a$ such that $\int_0^T a(t)dt = 1$, we can derive the desired equation first by:

$$(A.6) \implies S_T^x = \frac{1}{\sigma(t)} D_t S_T^x \implies \frac{1}{x} S_T^x = \frac{1}{x\sigma(t)} D_t S_T^x, \tag{A.8}$$

and then by using the function $a(t)$ to get the following equalities:

$$\frac{1}{x}S_T^x = \frac{1}{x}S_T^x \cdot 1 = \frac{1}{x}S_T^x \int_0^T a(t)dt = \int_0^T a(t)\frac{1}{x}S_T^x dt. \tag{A.9}$$

To conclude the proof of equation (A.7) we exchange $\frac{1}{x}S_T^x$ with the identity (A.8):

$$\frac{1}{x}S_T^x \overset{(A.9)}{=} \int_0^T a(t)\frac{1}{x}S_T^x dt \overset{(A.8)}{=} \int_0^T \frac{a(t)}{x\sigma(t)} D_t S_T^x dt.$$

∎

## A.2 Program Code

### A.2.1 Matlab Code

Program 1 - `bmotion.m` from [19]

```
function BM = bmotion()
% BMOTION — simulates a path of Brownian Motion
% Algorithm from Schoutens

N = 10^4;
Dt = 1/N;

BM = zeros(1,N);                    % Brownian motion
v = normrnd(0,1,[1 N]);

for n=2:N                           % BM(1) = 0
   BM(n) = BM(n−1) + sqrt(Dt)*v(n);
end

end
```

Program 2 - `compoisson.m` from [9]

```
function [CPP, times] = compoisson(intensity, var)
% COMPOISSON — Compound Poisson Process path with Gaussian jumps
% Extension of Poisson process—algorithm from Schoutens

T = 1;       % Interval: [0,T]
M = T*10^4; % Grid size (10.000 per unit)

N = poissrnd(T*intensity);  % Number of jumps
u = T*sort(rand(1,N));      % Jump times
times = round(M*u);         % Jump times for grid

J = cumsum(normrnd(0,var,[1 N]));  % N(0,var) jump sizes

CPP = zeros(1,M);           % Compound Poisson Process

CPP(1:times(1)) = 0;
for k=2:N
    CPP(times(k−1)+1:times(k))=J(k−1);
end
CPP(times(k)+1:M) = J(N);

end
```

Program 3 - `invgrnd.m` from [19]

```
function VEC = invgrnd(delta,gamma,N)
%INVGRND -  Draws Inverse Gaussian random variables
%           Algorithm from Schoutens
VEC = zeros(1,N);
v = randn(1,N);
y = v.^2;
x = (delta/gamma) + y./(2*gamma^2) -...
    sqrt(4*delta*gamma*y + y.^2)./(2*gamma^2);
u = rand(1,N);

for n=1:N
   if(u(n) <= delta/(delta + gamma*x(n)))
       VEC(n) = x(n);
   else
       VEC(n) = delta^2/(gamma^2*x(n));
   end
end
end
```

Program 4 - `NIGP.m` from [15]

```
function X = NIGP(alpha, beta, delta)
% NIGP - Simulates a path of the Normal Inverse Gaussian Process
% Algorithm from Korn
N = 10^4;
Dt = 1/N;

X = zeros(1,N+1);                % NIG-Process

gamma = sqrt(alpha^2-beta^2);
G = invgrnd(delta*Dt, gamma, N+1);
Y = normrnd(0,1, [1 N+1]);

X(1) = sqrt(G(1))*Y(1)+beta*G(1)*Dt;
for i=2:N+1
    X(i) = X(i-1) + sqrt(G(i))*Y(i)+beta*G(i)*Dt;
end

end
```

### A.2.2 Code for BNS Simulation

Supporting functions.

Program 5 - `Wplus.m`

```
function RES = Wplus(delta,gamma,X)
% WPLUS - Tail Mass Function for the
%        IG(delta, gamma) distribution

c = delta/sqrt(2*pi);
RES = c*(1./sqrt(X)).*exp(-0.5*gamma^2*X);

end
```

Program 6 - `Winv.m`

```
function res = Winv(x, X, WX)
% WINV - inverse of Wplus
%        The input WX is the result of running
%        WX = Wplus(delta,gamma,X)

I = find(WX<=x,1,'first');
res = X(I);

end
```

Program 7 - `ind.m`

```
function T = ind(test,lim)
% Supporting function: the indicator function for vectors.
T = test < lim;
end
```

Main function:

Program 8 - `mainBNS.m`

```
function [BNS, s2] = mainBNS(delta, gamma, inV, inP, r, rho, lam)
% MAINBNS -  Calls the other functions

[Z, fZ] = subord(delta, gamma, lam);
s2 = sigma2(fZ, lam, inV);      % inV : initial volatility
BNS = bns(delta,gamma,inP,r,rho,lam,s2,Z); % inP : initial price

end
```

Program 9 - subord.m

```
function [fZ,Z] = subord(delta,gamma,lambda)
%SUBORD — Simulates the subordinator Z_t and the Levy integral

M = 10^6;
dm = 1/M;
X = (0:dm:5);
WX = Wplus(delta,gamma,X);

% Finding trunctuating sum (slow)
NE = 0.5;
WinvNE = Winv(NE/lambda,X,WX);
while (WinvNE > dm)
   NE = NE+10;
   WinvNE = Winv(NE/lambda,X,WX);
end

tau = cumsum(exprnd(1, [1 2*NE]));

I = find(tau<NE);

N = length(I);
tau = tau(I);
u = rand(1,N);

T = 10000;
Z = zeros(1,T);
fZ = zeros(1,T);

% ——————————————— Algorithm based on Korn
S = zeros(1,N);
for i=1:N
    S(i) = Winv(tau(i)/lambda,X,WX);
end

for t=1:T
Sexp = zeros(1,N);
    for i=1:N
        Sexp(i) = S(i)*exp(u(i)*(t/T));
    end
    it = ind(u,t/T);
    Z(t) = dot(S,it);
    fZ(t) = dot(Sexp,it);
end

end
```

Program 10 - `sigma2.m`

```
function S2 = sigma2(fZ, lambda, init)
% SIGMA2 - Simulates the stochastic volatility for the BNS model.
% fZ is the result of running subord.m

T = 10^4;
S2 = zeros(1,T);
S2(1) = init; % sigma^2(0)>0.

for t=2:T
    S2(t) = exp(-lambda*(t/T))*(S2(1) + fZ(t));
end

end
```

Program 11 - `bns.m`

```
function BNSe = bns(delta,gamma,S0,r,rho,lambda,s2,Z)
% BNS -  Simulates the BNS model
% Z is the result of subord.m
% s2 is the result of sigma2.m

T = 10^4;
Dt = 1/T;
BNS = zeros(1,T);
BNS(1) = log(S0);

% Wiener Process
WP = sqrt(Dt)*normrnd(0,1,[1 T]);

% Interest rate minus the cumulant generating function
d = r - lambda*rho*delta/sqrt(gamma-2*rho);

for t=2:T
   BNS(t) = BNS(t-1) + (d - s2(t)/2)*Dt +...
                sqrt(s2(t))*WP(t) + rho*(Z(t)-Z(t-1));
end

BNSe = exp(BNS); % Exponential form.

end
```

Program 12 - `bnsWeights.m`

```
function [deltaW, gammaW] = bnsWeights(s2, BM, x)
%BNSWEIGHT - Calculates the Malliavin weights for the
%             BNS model. Assume T=1.

N = length(BM);
% Delta weight; calculated by Ito integral approximation
partsum = zeros(1,N-1);
s = sqrt(s2);
for i=1:N-1
   partsum(i) = (1/s(i))*(BM(i+1)-BM(i));
end
deltaW = (1/x)*sum(partsum);


% Gamma weight; calculated by the Trapezoidal rule.
s2div = 1./s2;
s2div(2:N-1) = 2*s2div(2:N-1);
gammaW = deltaW^2 - (1/x)*deltaW + (1/(2*N*x^2))*sum(s2div);

end
```

Program 13 - `IGproc.m` from [19]

```
function IP = IGproc(delta, gamma)
% IGPROC -  Simulates an IG(delta,gamma) Levy process
%           Algorithm from Schoutens

N = 10^4;
Dt = 1/N;

% Draw N-1 inverse Gaussian random numbers
IGrnd = invgrnd(delta*Dt, gamma, N-1);

IP = zeros(1,N);    %IP(1) = 0
for n=2:N
   IP(n) = IP(n-1) + IGrnd(n-1);
end

end
```

## A.3 Bibliography


[1] Barndorff-Nielsen, O.E., *Processes of normal inverse Gaussian type.* Finance and Stochastics, 2: 41-68, 1998.

[2] Barndorff-Nielsen, O.E., Shephard, N., *Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics.* J.R. Stat. Soc. Ser. B Stat. Methodol., 63 (2), 167-241, 2001.

[3] Benth, F.E., Groth, M., Kettler, P., *A Quasi-Monte Carlo Algorithm for the Normal Inverse Gaussian Distribution and valuation of financial derivatives.* International Journal of Theoretical and Applied Finance, 9(6):843-867, 2006.

[4] Benth, F.E., Groth, M., Wallin, O., *Derivative-free Greeks for the Barndorff-Nielsen and Shephard stochastic volatility model.* Stochastics. Vol. 82, No.3 291-313, 2010.

[5] Ben-Hamou, E., *Application of Malliavin Calculus and Wiener Chaos to Option Pricing Theory.* Ph.D Thesis, London School of Economics, 2000.

[6] Broadie, M., Glasserman, P., *Estimating security price derivatives using simulation*, Management Science, 24:169-285, 1996.

[7] Capiński, M., Kopp, E., *Measure, Integral and Probability.* Second Edition. Springer, 2004.

[8] Carmona, R., Tehranchi, M., *Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective.* Springer, 2006.

[9] Cont, R., Tankov, P., *Financial Modelling with Jump Processes.* Chapman & Hall, 2004.

[10] Davis, M., Johansson, M., *Malliavin Monte Carlo Greeks for jump diffusions.* Stochastic Process. Appl., 116(1):101-129. 2006.

[11] Di Nunno, G., Øksendal, B., Proske, F., *Malliavin Calculus for Lévy Processes with Applications to Finance*. Springer, 2008.

[12] Jäckel, P., *Monte Carlo methods in finance.* John Wiley & Sons Ltd, 2002.

[13] Fournié, E., et al., *Applications of Malliavin Calculus to Monte Carlo methods in finance.* Finance Stoch. 3, 391-412, 1999.

[14] Fournié, E., et al., *Applications of Malliavin Calculus to Monte Carlo methods in finance. II.* Finance Stoch. 5, 201-236, 2001.

[15] Korn, R., Korn, E., Kroisandt, G., *Monte Carlo Methods and Models in Finance and Insurance.* Chapman & Hall, 2010.

[16] Raible, S., *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts.* Ph.D Thesis, University of Freiburg, 2000.

[17] Schachermayer, W., Teichmann, J., *How close are the option pricing formulas of Bachelier and Black-Merton-Scholes?* Preprint, 2005.

[18] Schilling, R. *Measures, Integrals and Martingales.* Cambridge University Press, 2005.

[19] Schoutens, W., *Lévy Processes in Finance.*, John Wiley & Sons Ltd, 2003.

[20] Wilmott, P., *Paul Wilmott introduces Quantitative Finance.* John Wiley & Sons Ltd, 2007.

[21] Øksendal, B., *Stochastic Differential Equations.* 6th Edition. Springer, 2007.